Prediction Markets

Theory and applications

Edited by

Leighton Vaughan Williams

ROUTLEDGE

Routledge International Studies in Money and Banking

ISBN 978-0-415-57286-6
9 780415 572866

Routledge
Taylor & Francis Group

Prediction Markets

Theory and applications

Edited by Leighton Vaughan Williams

Routledge
Taylor & Francis Group
LONDON AND NEW YORK

First published 2011
by Routledge
2 Park Square, Milton Park, Abingdon, Oxon OX14 4RN

Simultaneously published in the USA and Canada
by Routledge
711 Third Avenue, New York, NY 10017

Routledge is an imprint of the Taylor & Francis Group, an informa business

British Library Cataloguing in Publication Data
A catalogue record for this book is available from the British Library

Library of Congress Cataloging in Publication Data
Prediction markets : theory and applications / edited by Leighton Vaughan Williams.
p. cm.
Includes bibliographical references and index.
1. Business forecasting. 2. Economic forecasting. 3. Forecasting.
I. Vaughan Williams, Leighton.
HD30.27.P74 2011
303.49–dc22

2010049035

ISBN: 978-0-415-57286-6 (hbk)
ISBN: 978-0-203-81552-6 (ebk)

Typeset in Times
by Wearset Ltd, Boldon, Tyne and Wear

MIX
Paper from responsible sources
FSC® C004839

Printed and bound in Great Britain by
TJI Digital, Padstow, Cornwall

Prediction Markets

How can we effectively aggregate disparate pieces of information that are spread among many different individuals? In other words, how can we best access the 'wisdom of the crowd'? Prediction markets, which are essentially speculative markets created for the purpose of aggregating information and making predictions, offer the answer to this question. The effective use of these markets has the potential not only to help forecast future events on a national and international level, but also to assist companies in providing, for example, improved estimates of the potential market size for a new product idea or the launch date of new products and services.

The markets have already been used to forecast uncertain outcomes ranging from influenza to the spread of infectious diseases, to the demand for hospital services, to the box office success of movies, climate change, vote shares and election outcomes, to the probability of meeting project deadlines. The insights gained also have many potentially valuable applications for public policy more generally. These markets offer substantial promise as a tool of information aggregation as well as forecasting, whether alone or as a supplement to other mechanisms like surveys, group deliberations and expert opinion. Moreover, they can be applied at a macroeconomic and microeconomic level to yield information that is valuable for government and commercial policy-makers and which can be used for a number of social purposes.

This volume of original readings, contributed by many of the leading experts in the field, marks a significant addition to the base of knowledge about this fascinating subject area. The book should appeal to all those with an interest in economics, forecasting or public policy, and in particular those with an interest in the study of money, investment and risk.

Leighton Vaughan Williams is Professor of Economics and Finance and Director of the Betting Research and the Political Forecasting Unit at Nottingham Business School, Nottingham Trent University, UK.

Routledge international studies in money and banking

1 **Private Banking in Europe**
Lynn Bicker

2 **Bank Deregulation and Monetary Order**
George Selgin

3 **Money in Islam**
A study in Islamic political economy
Masudul Alam Choudhury

4 **The Future of European Financial Centres**
Kirsten Bindemann

5 **Payment Systems in Global Perspective**
Maxwell J Fry, Isaak Kilato, Sandra Roger, Krzysztof Senderowicz, David Sheppard, Francisco Solis and John Trundle

6 **What is Money?**
John Smithin

7 **Finance**
A characteristics approach
Edited by David Blake

8 **Organisational Change and Retail Finance**
An ethnographic perspective
Richard Harper, Dave Randall and Mark Rouncefield

9 **The History of the Bundesbank**
Lessons for the European Central Bank
Jakob de Haan

10 **The Euro**
A challenge and opportunity for financial markets
Published on behalf of Société Universitaire Européenne de Recherches Financières (SUERF)
Edited by Michael Artis, Axel Weber and Elizabeth Hennessy

11 **Central Banking in Eastern Europe**
Edited by Nigel Healey and Barry Harrison

12 **Money, Credit and Prices Stability**
Paul Dalziel

13 **Monetary Policy, Capital Flows and Exchange Rates**
Essays in memory of Maxwell Fry
Edited by William Allen and David Dickinson

14 **Adapting to Financial Globalisation**
Published on behalf of Société Universitaire Européenne de Recherches Financières (SUERF)
Edited by Morten Balling, Eduard H. Hochreiter and Elizabeth Hennessy

15 **Monetary Macroeconomics**
A new approach
Alvaro Cencini

16 **Monetary Stability in Europe**
Stefan Collignon

17 **Technology and Finance**
Challenges for financial markets, business strategies and policy makers
Published on behalf of Société Universitaire Européenne de Recherches Financières (SUERF)
Edited by Morten Balling, Frank Lierman, and Andrew Mullineux

18 **Monetary Unions**
Theory, history, public choice
Edited by Forrest H. Capie and Geoffrey E Wood

19 **HRM and Occupational Health and Safety**
Carol Boyd

20 **Central Banking Systems Compared**
The ECB, the pre-euro Bundesbank and the Federal Reserve System
Emmanuel Apel

21 **A History of Monetary Unions**
John Chown

22 **Dollarization**
Lessons from Europe and the Americas
Edited by Louis-Philippe Rochon and Mario Seccareccia

23 **Islamic Economics and Finance: A Glossary, 2nd Edition**
Muhammad Akram Khan

24 **Financial Market Risk**
Measurement and analysis
Cornelis A. Los

25 **Financial Geography**
A Banker's view
Risto Laulajainen

26 **Money Doctors**
The experience of international financial advising 1850–2000
Edited by Marc Flandreau

27 **Exchange Rate Dynamics**
A new open economy macroeconomics perspective
Edited by Jean-Oliver Hairault and Thepthida Sopraseuth

28 **Fixing Financial Crises in the 21st Century**
Edited by Andrew G. Haldane

29 **Monetary Policy and Unemployment**
The U.S., euro-area and Japan
Edited by Willi Semmler

30 **Exchange Rates, Capital Flows and Policy**
Edited by Peter Sinclair, Rebecca Driver and Christoph Thoenissen

31 **Great Architects of International Finance**
The Bretton Woods era
Anthony M. Endres

32 **The Means to Prosperity**
Fiscal policy reconsidered
Edited by Per Gunnar Berglund and Matias Vernengo

33 **Competition and Profitability in European Financial Services**
Strategic, systemic and policy issues
Edited by Morten Balling, Frank Lierman and Andy Mullineux

34 **Tax Systems and Tax Reforms in South and East Asia**
Edited by Luigi Bernardi, Angela Fraschini and Parthasarathi Shome

35 **Institutional Change in the Payments System and Monetary Policy**
Edited by Stefan W. Schmitz and Geoffrey E. Wood

36 **The Lender of Last Resort**
Edited by F.H. Capie and G.E. Wood

37 **The Structure of Financial Regulation**
Edited by David G. Mayes and Geoffrey E. Wood

38 **Monetary Policy in Central Europe**
Miroslav Beblavý

39 **Money and Payments in Theory and Practice**
Sergio Rossi

40 **Open Market Operations and Financial Markets**
Edited by David G. Mayes and Jan Toporowski

41 **Banking in Central and Eastern Europe 1980–2006**
A comprehensive analysis of banking sector transformation in the former Soviet Union, Czechoslovakia, East Germany, Yugoslavia, Belarus, Bulgaria, Croatia, the Czech Republic, Hungary, Kazakhstan, Poland, Romania, the Russian Federation, Serbia and Montenegro, Slovakia, Ukraine and Uzbekistan.
Stephan Barisitz

42 **Debt, Risk and Liquidity in Futures Markets**
Edited by Barry A. Goss

43 **The Future of Payment Systems**
Edited by Stephen Millard, Andrew G. Haldane and Victoria Saporta

44 **Credit and Collateral**
Vania Sena

45 **Tax Systems and Tax Reforms in Latin America**
Edited by Luigi Bernardi, Alberto Barreix, Anna Marenzi and Paola Profeta

46 **The Dynamics of Organizational Collapse**
The case of Barings Bank
Helga Drummond

47 **International Financial Co-operation**
Political economics of compliance with the 1988 Basel Accord
Bryce Quillin

48 **Bank Performance**
A theoretical and empirical framework for the analysis of profitability, competition and efficiency
Jacob Bikker and Jaap W.B. Bos

49 **Monetary Growth Theory**
Money, interest, prices, capital, knowledge and economic structure over time and space
Wei-Bin Zhang

50 **Money, Uncertainty and Time**
Giuseppe Fontana

51 **Central Banking, Asset Prices and Financial Fragility**
Éric Tymoigne

52 **Financial Markets and the Macroeconomy**
Willi Semmler, Peter Flaschel, Carl Chiarella and Reiner Franke

53 **Inflation Theory in Economics**
Welfare, velocity, growth and business cycles
Max Gillman

54 **Monetary Policy Over Fifty Years**
Heinz Herrman (Deutsche Bundesbank)

55 **Designing Central Banks**
David Mayes and Geoffrey Wood

56 **Inflation Expectations**
Peter J.N. Sinclair

57 **The New International Monetary System**
Essays in honour of Alexander Swoboda
Edited by Charles Wyplosz

58 **Taxation and Gender Equity**
A comparative analysis of direct and indirect taxes in developing and developed countries
Edited by Caren Grown and Imraan Valodia

59 **Developing Alternative Frameworks for Explaining Tax Compliance**
Edited by James Alm, Jorge Martinez-Vazquez and Benno Torgler

60 **International Tax Coordination**
An interdisciplinary perspective on virtues and pitfalls
Edited by Martin Zagler

61 **The Capital Needs of Central Banks**
Edited by Sue Milton and Peter Sinclair

62 **Monetary and Banking History**
Edited by Geoffrey E. Wood, Terence Mills and Nicholas Crafts

63 **New Approaches to Monetary Economics and Theory**
Interdisciplinary perspectives
Edited by Heiner Ganßmann

64 **Social Banks and the Future of Sustainable Finance**
Edited by Olaf Weber and Sven Remer

65 **Policy Makers on Policy**
The Mais lectures
Edited by Forrest H. Capie and Geoffrey E. Wood

66 **Prediction Markets**
Theory and applications
Edited by Leighton Vaughan Williams

Contributors

Bin-Tzong Chie is Assistant Professor in the Department of Industrial Economics of the Tamkang University. He received his PhD in Economics at the National Chengchi University in 2007. He is a Research Fellow at AI-Econ Research Center, National Chengchi University. His research interests include agent-based computational economics, experimental economics and decision-making under uncertainty.

Richard Borghesi earned a PhD from the University of Florida in 2004. His research interests include corporate finance, corruption, market efficiency and prediction markets. He has published in *The Journal of Prediction Markets*, *Financial Management*, *Southern Economic Journal*, *Finance Research Letters*, *Journal of Applied Business Research*, *Journal of Economics and Business*, and *Applied Financial Economics*.

Chung-Ching Tai has been Assistant Professor of Economics since 2009 at Tunghai University, Taiwan. He has a PhD from the National Chengchi University, Taiwan. His research focuses on agent-based computational economics and experimental economics. His current research topics include prediction markets, agent-based double-auction markets and cognitive market experiments.

Chen-yuan Tung is Professor at the Graduate Institute of Development Studies and Director of the Center for Prediction Markets, National Chengchi University, Taiwan. He received his PhD majoring in international affairs from the School of Advanced International Studies, Johns Hopkins University. His expertise focuses on international political economy, Chinese economic development, Taiwan–United States–China trilateral relations and prediction markets.

Les Coleman teaches Finance at the University of Melbourne. Les trained originally as an engineer, and his management PhD thesis was published as *Why Managers and Companies Take Risks* (Springer, 2006). His principal research objective is to build two bodies of theory on the bases of financial decisions, and what leads to uncertainty and loss of value by corporations with a preliminary version of the latter in his latest book, *Risk Strategies:*

Dialling up Optimum Firm Risk (Gower, 2009). His papers have been published in numerous journals, he is a regular contributor to the media and delivers executive education programmes in Australia and overseas, and has received research and teaching awards.

Karen Croxson holds a PhD in economics from the University of Oxford. In 2007 she was awarded the Rank-Manning Junior Research Fellowship at New College and began working on market microstructure and prediction markets. In 2008 she joined the Oxford-Man Institute of Quantitative Finance. Some of her recent research has exploited high-frequency data from prediction markets. Other strands have addressed financial regulation, digital piracy, leadership and coordination. Since contributing to this volume, Karen has joined the McKinsey Global Institute.

Andreas Graefe is Research Associate at the Karlsruhe Institute of Technology, where he received his PhD in Economics. His research focus is on the development of forecasting methods that are simple to use, easy to understand, and thus can aid decision-making in the real world. He is also the prediction markets editor of *Foresight – The International Journal of Applied Forecasting.*

Johnnie Johnson is Professor of Decision and Risk Analysis and Director of the Centre for Risk Research in the School of Management at the University of Southampton. He has published widely in the areas of risk taking and decision making under uncertainty, with particular reference to decisions made in speculative markets. Johnnie has a particular interest in developing new approaches for forecasting outcomes in prediction and betting markets.

Iván Marinovic is currently a PhD student at the department of Accounting Information and Management at Kellogg, where he is expected to complete his candidacy shortly. He joined Northwestern after earning an MSc in Economic Theory from the University of Toulouse in France. His current research focuses on financial reporting and disclosure, the efficiency of information systems and the economics of science.

Charles Noussair is Professor of Economics at Tilburg University, and an Adjunct Faculty Member at Emory University and the University of Melbourne. He has a BA from the University of Pennsylvania and an MS and PhD from the California Institute of Technology. His research applies the methods of experimental economics to a variety of topics in micro-and macroeconomics. His research has been published in the *American Economic Review*, the *Journal of Political Economy*, *Econometrica*, the *Journal of Finance*, and the *Economic Journal.*

Anastasios Oikonomidis is undertaking his PhD at the University of Southampton, where he has held the post of Research Assistant in the Centre for Risk Research. His primary current research interest concerns the investigation of information efficiency in football betting markets.

Marco Ottaviani has taught at Northwestern since 2007. He received a PhD in Economics from the Massachusetts Institute of Technology in 1996. His research focuses on information aggregation in markets and organisations. He is currently working on the provision of incentives for information intermediaries, the development of prediction markets to improve corporate and public decision-making, and the rationale for government policies intended at mandating information disclosure and protecting consumers. He is managing editor of the *Review of Economic Studies* and associate editor of the *B.E. Journal in Theoretical Economics* and *The Journal of Prediction Markets*.

Peter Norman Sørensen has been affiliated with the University of Copenhagen since 1998. He received a PhD in Economics from the Massachusetts Institute of Technology in 1996. His research focuses on social learning, strategic communication and information aggregation. He is currently working on the design of prediction markets and on the liquidity effects of central bank interventions. He is editor of *Economica* and associate editor of *Macroeconomic Dynamics*. He is also head of the Finance Research Unit in Copenhagen.

Lionel Page is currently affiliated with the Queensland University of Technology and the University of Cambridge. His research interests include experimental and behavioural economics, applied econometrics, the economics of education and decision making under risk and uncertainty.

Rodney J. Paul is currently Professor of Economics in the Department of Finance at St Bonaventure University. He received his PhD in Applied Economics from Clemson University in 2000. His work in the prediction markets field involves a variety of studies based on prices formed in sports wagering markets. He has studied market efficiency across a wide range of sports, investigated sportsbook behaviour and behavioural biases of bettors through actual sportsbook betting percentages, and has used prices formed in these markets as a way of understanding other sports-related topics. Other research interests include attendance studies and time-series macroeconomics.

John Peirson is Senior Lecturer in Economics at the University of Kent. He has a BA in Mathematics and Economics from the University of Cambridge and a PhD in Economics from Southampton University. He has worked at the Universities of Sheffield, York and Kent, and worked for the UK government. He has published widely in the fields of energy economics, the economics of uncertainty and gambling, transport economics and new economic geography. He has advised the UK, Nigerian and Philippine governments and the European Union on energy and transport economics.

Adi Schnytzer has been an Associate Professor of Economics in the Department of Economics at Bar Ilan University since 1997. He received his BSc and MEc from Monash University and his DPhil from Oxford University. His research focuses on sportometrics, comparative economics and the economics of prediction and betting markets. He currently serves on the Editorial Board of *The Journal of Gambling Business and Economics*.

Shu G. Wang is Associate Professor in the Department of Economics and also serves as the Associate Director of the AI-ECON Research Center, National Chengchi University. Dr Wang holds a PhD in Economics from Purdue University. His research interests are mainly microeconomics, institutional economics, law and economics and, recently, agent-based computational economics and experimental economics.

Shu-Heng Chen is a Distinguished Professor in the Department of Economics and Director of the AI-ECON Research Center at National Chengchi University. He also serves as Vice Chair of the IEEE Computational Finance & Economics Technical Committee, and as editor-in-chief of the *Journal of New Mathematics and Natural Computation*, associate editor of the *Journal of Economic Behaviour and Organization* and editor of the *Journal of Economic Interaction and Coordination*. Professor Chen holds a PhD in Economics from the University of California at Los Angeles. He has published more than 150 publications in journals and volumes.

Bernd Skiera is a Professor of Electronic Commerce at the University of Frankfurt, Germany, and a member of the board of the E-Finance Lab at the House of Finance in Frankfurt. His research focuses on prediction markets, the impact of information technology on the financial service industry, online marketing, pricing and customer management. His work has been published in *Management Science*, *Marketing Science*, *Journal of Marketing Research*, *Journal of Marketing*, *Economic Letters*, *Journal of Product Innovation Management*, *Journal of Interactive Marketing*, and the *European Journal of Operational Research*.

Michael A. Smith is a Senior Lecturer in Economics at Leeds Business School. His PhD was awarded by Nottingham Trent University, based on an empirical study of information efficiency in a range of betting markets. His current research interests centre on the structure and dynamics of betting exchange and bookmaker markets.

Erik Snowberg is Assistant Professor of Economics and Political Science at the California Institute of Technology. Erik's research is at the intersection of economics and politics, examining how the economy affects elections, and how elections affect the economy. He holds a few undergraduate degrees from the Massachusetts Institute of Technology and a PhD from the Stanford Graduate School of Business.

Martin Spann is a Professor of Electronic Commerce at the School of Management of the Ludwig-Maximilians-University in Munich, Germany. He received his PhD from Goethe University in Frankfurt, Germany. Martin's current research interests are electronic commerce, pricing, auctions, innovation management, prediction markets and social network analysis. Martin's work has been published in *Marketing Science, Management Science, Information Systems Research, MIS Quarterly, Journal of Marketing, Journal of Product Innovation Management, Journal of Interactive Marketing, Journal of Forecasting, European Journal of Operational Research* and other journals.

Tzu-Chuan Chou received MS and PhD degrees from the Department of Computer Science and Information Engineering, Tamkang University, Taiwan, in 1998 and 2004, respectively. He was a postdoctoral fellow in the Institute of Information Science, Academia Sinica, Taiwan from 2004 to 2009. He is currently a postdoctoral fellow at the Graduate Institute of Development Studies, National Chengchi University, Taiwan. His research interests include prediction markets, clustering algorithms, information retrieval and image compression.

Leighton Vaughan Williams is Professor of Economics and Finance, Head of Economics Research and Director of the Betting Research Unit and of the Political Forecasting Unit at Nottingham Business School, Nottingham Trent University. He advises the UK government on the taxation and regulation of betting and gaming, and has published extensively in the fields of risk, asymmetric information, financial, betting and prediction markets. He is editor of *Economic Issues, The Journal of Gambling Business and Economics* and the *Journal of Prediction Markets.*

Andrew P. Weinbach has been teaching in the Wall College of Business at Coastal Carolina University, where he has held the Colonel Lindsey H. Vereen Endowed Professorship since 2006. His research focuses on the determinants of consumer interest in live sporting events, including forecasting of television viewing, bettor participation and live event attendance. Prior to receiving his PhD from Clemson University, Weinbach worked in the field of information technology as a network engineer, consultant and project manager, specialising in large-scale technology deployment projects for Fortune 500/Global 500 companies.

Justin Wolfers is Associate Professor of Business and Public Policy at the Wharton School of the University of Pennsylvania and a Research Associate of the National Bureau of Economic Research. His research interests include labour economics, macroeconomics, political economy, economics of the family, social policy, law and economics, public economics and behavioural economics.

Yiling Chen is an Assistant Professor of computer science at Harvard University. She has a BS from Renmin University of China, an MS from Tsinghua University, Beijing, China, and a PhD from the Pennsylvania State University. Prior to her appointment at Harvard, she worked at Yahoo! Research, New York. Her research lies at the intersection of computer science and economics. She is interested in designing and analysing social and organisational systems according to both computational and economic objectives.

Eric Zitzewitz is Associate Professor of Economics at Dartmouth College. His research interests include the financial services industry and market-based forecasting and he has a special interest in forensic economics. He received a PhD in economics from MIT and an AB in economics from Harvard College. He has studied stake pricing and late trading in mutual funds, the economic impact of the Iraq War, the effects of fair disclosure regulation and bias and related distortions from equity analysts, financial journalists and Olympic judges.

Contents

1 Introduction

Leighton Vaughan Williams

How can we effectively aggregate disparate pieces of information that are spread among many different individuals? In other words, how can we best access the 'wisdom of the crowd'? Prediction markets, which are speculative markets created for the purpose of aggregating information and making predictions, are a means of addressing this issue. Their theoretical underpinning derives from the efficient-markets hypothesis and stems from the view that relevant information concerning the likelihood of future events is dispersed among the opinions and intuitions of many people. While the mechanisms underlying prediction markets vary, they all offer a means of aggregating this information. Many of these markets are open to the public, while others are open to particular groups.

Prediction markets can be used to provide forecasts of the probability and the mean and median outcomes of future events, as well as correlations between these events. They also have many potentially valuable applications for public policy (Paton *et al.*, 2009, 2010).

The markets have already been used to forecast uncertain outcomes ranging from influenza outbreaks (Wang *et al.*, 2009) to the spread of infectious diseases (Polgreen *et al.*, 2007), the demand for hospital services (Rajakovich and Vladimirov, 2009) the box office success of movies, climate change, vote shares, election outcomes (Rhode and Strumpf, 2004; Wolfers and Zitzewitz, 2004; Snowberg *et al.*, 2005) and the probability of meeting project deadlines at Google (Leigh and Wolfers, 2007). Prediction markets may also be used as a mechanism to help market participants hedge their exposure to risk (Athanasoulis *et al.*, 1999).

Even so, some researchers have questioned how far prediction markets are able to outperform other means of forecasting (Erikson and Wlezien, 2008). It is also suggested that they may be open to manipulation (Wolfers and Leigh, 2002), though this might actually aid prediction market accuracy (Hanson and Oprea, 2009). Again, they may not provide efficient forecasts of low probability events (Smith *et al.*, 2006; Wolfers and Zitzewitz, 2004), and may be open to systematic biases, such as optimism bias (Cowgill *et al.*, 2009) and the favourite-longshot bias (Vaughan Williams and Paton, 1997).

The effective use of prediction markets has the potential, however, not only to help forecast events at a national and international level, but also to assist companies in, for example, providing improved estimates of the potential market size for a new product idea or the launch date of new products and services. Examples of companies that have used internal prediction markets for a range of business forecasts include Hewlett-Packard (Chen and Plott, 2002), Google (Cowgill *et al.*, 2009) and General Electric (Spears *et al.*, 2009). The success and potential of these markets in predicting public events and corporate outcomes has therefore generated substantial interest among social scientists, policymakers and the business community. The insights gained also have many potentially valuable applications for policy more generally, not least when accurate forecasts are required in relation to quantifiable targets. Moreover, the information provided by prediction markets will have value in the advance warning managers may be given of weak performance in identifiable areas. This can help improve resource allocation.

Important research questions include the impact of prediction market design choices on performance (Spann and Skiera, 2003) and the impact of the nature of rewards on the level of accuracy of prediction markets (Servan-Schreiber *et al.*, 2004). Indeed, the design of the incentive programme may be critical to optimising performance, insofar as people may invest more thought and energy into expressing their opinion when there is a meaningful incentive to do so.

Overall, the balance of opinion provided by previous research suggests that well-designed prediction markets can offer substantial promise as a tool of information aggregation and forecasting, whether alone or as a supplement to other mechanisms like surveys, group deliberations and expert opinion. Moreover, they can be applied at a macroeconomic and microeconomic level to yield information that is valuable for government and commercial policymakers, and that can be used for a number of social purposes.

This volume of original readings marks a significant addition to the base of knowledge about this fascinating subject area. What is provided is a collection of readings that draws on the expertise of many of the leading contributors in the field. The chapters are not only novel and original, but also set the subject within the existing framework of literature. As such, this book should serve as a valuable asset for those who are coming fresh to the subject, as well as for those who are more familiar with the subject matter. The contributors hail from a host of prestigious institutions located as far afield as Australia, Denmark, Germany, the Netherlands, Israel, Taiwan, the United Kingdom and the United States.

In many cases, the contributions would, in my opinion, have gone on to be published in top-ranked journals, but the authors lent their support instead to the idea of a single volume that would help promote this field of research to a wider audience. In all cases, the authors have provided contributions that are valuable and important, and which contribute something significant to help meet the burgeoning growth of interest in the theory and applications of prediction markets. It has been a pleasure to edit this book, and my deepest gratitude goes to all involved.

References

Athanasoulis, S., Shiller, R. and Rietz, T. (1999), *Macro-Markets and Financial Security, Economic Policy Review*, 5, 1, 21–39.

Chen, K.-Y. and Plott, C.R. (2002), *Information Aggregation Mechanisms: Concept, Design and Implementation for a Sales Forecasting Problem*, CalTech, Division of Humanities and Social Sciences, Working Paper 1131.

Cowgill, B., Wolfers, J. and Zitzewitz, E. (2009), *Using Prediction Markets to Track Information Flows: Evidence from Google*, Working Paper, available at: www.bocowgill.com/GooglePredictionMarketPaper.pdf.

Erikson, R.S. and Wlezien, C. (2008), Are Political Markets Really Superior to Polls as Election Predictors? *Public Opinion Quarterly*, 72, 2, 190–215.

Hanson, R. and Oprea, R. (2009), A Manipulator Can Aid Prediction Market Accuracy, *Economica*, 76, 302, 304–311.

Leigh, A. and Wolfers, J. (2007), Prediction Markets for Business and Public Policy, *Melbourne Review*, 3,1, 7–15.

Paton, D., Siegel, D. and Vaughan Williams, L. (2009), The Growth of Gambling and Prediction Markets: Economic and Financial Implications, *Economica*, 76, 302, 219–224.

Paton, D., Siegel, D. and Vaughan Williams, L. (2010), Gambling, Prediction Markets and Public Policy, *Southern Economic Journal*, 76, 4, 878–883.

Polgreen, P.M., Nelson, F.D. and Newmann, G.R. (2007), Use of Prediction Markets to Forecast Infectious Disease Activity, *Clinical Infectious Diseases*, 44, 272–279.

Rajakovich, D. and Vladimirov, V. (2009), Prediction Markets as a Medical Forecasting Tool: Demand for Hospital Services, *The Journal of Prediction Markets*, 3, 2, 78–106.

Rhode, P.W. and Strumpf, K.S. (2004), Historical Presidential Betting Markets, *Journal of Economic Perspectives*, 18, 2, 127–141.

Servan-Schreiber, E.J., Wolfers, J., Pennock, D. and Galebach, A. (2004), Prediction Markets: Does Money Matter, *Electronic Markets*, 14, 243–251.

Smith, M., Paton, D. and Vaughan Williams, L. (2006), Market Efficiency in Person-to-Person Betting Exchanges, *Economica*, 73, 292, 673–689.

Smith, M., Paton, D. and Vaughan Williams, L. (2009), Do Bookmakers Possess Superior Skills to Bettors in Predicting Outcomes?, *Journal of Economic Behavior and Organization*, 71, 2, 539–549.

Snowberg, E., Wolfers, J. and Zitzewitz, E. (2005), Information (In)Efficiency in Prediction Markets, in: *Information Efficiency in Financial and Betting Markets*, ed. Vaughan Williams, L., Cambridge University Press, 366–386.

Spann, M. and Skiera, B. (2003), Internet-Based Virtual Stock Markets for Business Forecasting, *Management Science*, 49, 10, 1310–1326.

Spears, B., LaComb, C., Interrante, J., Barnett, J. and Senturk-Dogonaskoy, D. (2009), Examining Trader Behavior in Idea Markets: An Implementation of GE's Imagination Markets, *The Journal of Prediction Markets*, 3, 1, 17–39.

Vaughan Williams, L. and Paton, D. (1997), Why is there a Favourite-Longshot Bias in British Racetrack Betting Markets? *Economic Journal*, 107, 150–158.

Wang, S.-C., Tseng, J.-J., Li, S.-P. and Chen, S.-H. (2009), Prediction of Bird Flu A (H5N1) outbreaks in Taiwan by Online Auction: Experimental Results, *New Mathematics and Natural Computation*, 2, 271–280.

Wolfers, J. and Leigh, A. (2002), Three Tools for Forecasting Federal Elections: Lessons from 2001, *Australian Journal of Political Science*, 37, 2, 223–240.

Wolfers, J. and Zitzewitz, E. (2004), Prediction Markets, *Journal of Economic Perspectives*, 18, 107–126.

2 Modeling idea markets

Between beauty contests and prediction markets

Iván Marinovic, Marco Ottaviani and Peter Norman Sørensen

1 Introduction

This chapter is motivated by a recent surge of interest in evaluating ideas (such as new products or corporate projects) by means of trading mechanisms known as *idea markets*. In idea markets, ideas are associated to assets that are then traded among market participants who possess relevant information. In a baseline design, traders are rewarded on the basis of the value of their portfolio when the market is closed.[1] Idea markets hold great promise to initially complement and eventually substitute traditional techniques for market research, which is mostly based on costly surveys. In this chapter, we investigate the information aggregation properties of idea markets, focusing on the role played by the incentive to buy popular assets, and on how this incentive depends on the size of the market.

Proposers of idea markets are inspired by the performance of prediction markets, such as Iowa electronic markets, as well as horse-race betting markets.[2] A defining feature of prediction markets is that assets are liquidated on the basis of the ex-post realization of the underlying variable. Thus, traders in prediction markets have an incentive to maximize the value of their portfolio by picking the assets with the most favorable expected realizations of the underlying variable. The resulting equilibrium price in a prediction market should then reflect the traders' overall information, much like in a financial market driven by fundamentals.[3]

While participants in prediction markets gain from predicting accurately the final outcome, in idea markets participants are rewarded on their ability to accurately predict the choices of other market participants.[4] This is often because the ex-post performance of the underlying ideas is difficult to observe or verify.[5] To wit, the profitability of the ideas that are not implemented is simply not observed. Even when information about the profitability of an implemented idea is observed, typically this is only with sizeable noise and major delay. Without the ex-post validation that is characteristic of prediction markets, pure idea markets become beauty contests.

This chapter investigates how much information is aggregated when agents thus face mixed rewards based partly on ex-post realizations (as is typical of prediction markets) and partly on the most popular opinion (as in a beauty contest).

As a first attempt to formally model this tension, we analyze a variant of a model proposed by Morris and Shin (2002). The model features a set of agents endowed with both private and common information who simultaneously submit predictions. With some probability, agents are rewarded on the basis of the accuracy of their individual prediction relative to the final outcome. With complementary probability, agents are rewarded on the distance between their prediction and the average prediction across agents (the *consensus forecast*). This probabilistic interpretation of the rewards captures the notion that the market designer may only observe the value of the idea with some probability, which we refer to as the *prediction market intensity*.

The mixed nature of rewards means that agents care not only about their own assessment of the final outcome, but also – because of the incentives to coordinate with other agents – about their assessments of other agents' assessments, and about the assessments of other agents about the assessments of other agents, and so on. Higher order beliefs play a key role in these markets, as in Keynes' (1936) celebrated metaphor of financial markets as beauty contests.[6]

We examine how these beauty contest incentives affect the informativeness of the consensus forecast. We begin by considering pure markets, where agents only care about one type of incentive. A pure idea market is a beauty contest in which agents are only driven by the incentive to second guess other agents' beliefs so as to meet the consensus forecast. As a result, agents rely exclusively on common information, which is the most effective way to coordinate with other agents. But precisely because agents disregard private information, the informativeness of the consensus forecast in a beauty contest becomes very poor.[7]

In the opposite extreme case, a pure prediction market, the consensus forecast aggregates information much more effectively. An agent's optimal strategy consists of submitting a prediction that coincides with the agent's best predictor of the idea's value, which results in a consensus forecast that, at least to some extent, reflects agents' private information. Aggregating the private information, the consensus forecast is more accurate the greater the number of agents in the prediction market. However, the consensus forecast still overweights the common information relative to an optimal aggregate of all private and public information. As a result, the informativeness of the consensus forecast need not increase in the precision of common information. An increase in the precision of common information has two effects. Holding agents' strategies constant, more precise common information leads naturally to a more precise consensus forecast. These informational benefits, though, may be offset by agents' reactions: a more precise common signal induces agents to increase the weight they attach to the common signal at the expense of the private signal. Since the consensus forecast assigns excessive weight to common information, this effect leads to a redundancy of common information in the consensus forecast, eventually reducing its informativeness.

After reviewing the two pure cases, we examine the case in which agents have mixed incentives. Intuitively, in equilibrium, agents' predictions are a weighted average of their private and common signals. First and foremost, we establish that the informativeness of the consensus forecast increases in the intensity of prediction markets. By reducing agents' coordination concerns, a greater intensity of prediction markets lowers the *weight of common information* in agents' predictions, thereby reducing the redundancy of common information that affects the consensus forecast.

What is the impact of an increase in the number of (privately informed) agents on the informativeness of the consensus forecast? Our main result is a characterization of situations in which an increase in the number of agents has a detrimental effect on the informativeness of the consensus forecast. As we show, increasing the number of agents has two effects. On the one hand, there is the statistical benefit of a larger sample – more agents allow the consensus forecast to filter more efficiently the noise of agents' private signals. On the other hand, a larger sample also has an indirect effect: the distortions that arise from agents' *coordination incentives are worsened when there is a larger number of informed* agents. In a smaller sample, the influence of an individual agent over the consensus forecast is larger, so each agent is less concerned with second guessing the predictions of other agents. As the number of agents grows, the impact of any given agent on the consensus forecast decreases. Coordination becomes a more important concern for an agent. Agents' greater emphasis on coordination lowers the informativeness of the consensus forecast, even to the point of offsetting the statistical benefits of a larger sample.

We also find, in the general case, that better common information may deteriorate the informativeness of the consensus forecast.[8] The mechanism is similar to that described for a pure prediction market, but aggravated here by agents' coordination incentives. An increase in the quality of common information not only helps agents to make better predictions, but also helps them to coordinate more efficiently. Agents respond by relying more heavily on common information when they form their predictions. Consequently, the correlation between the consensus forecast and agents' private signals becomes weaker, particularly when the number of agents is large. This, in turn, lowers the informational content of the consensus when private signals are more accurate than common information.

Other theoretical work on prediction markets addresses different problems. Hanson and Oprea (2009) and Hahn and Tetlock (2007) analyze the incentives of prediction market agents to affect the market price in order to affect the decision based on that price. Ottaviani and Sørensen (2007) investigate the problem of outcome manipulation whereby agents affect the outcome on which the prediction market payoff is paid. Lieli and Nieto-Barthaburu (2009) analyze the effect of feedback on the operation of prediction markets when a decision-maker's intervention, based on the information revealed by the market, affects the probability of the underlying event.

We proceed by describing the model in Section 2. In Section 3 we introduce a notion of informativeness. In Section 4 we characterize the equilibrium and discuss its informational properties.

2 Model

The model is a variant of Morris and Shin (2002), with a finite number of agents rather than a continuum of agents.

A market designer is interested in learning the value of an idea, which is captured by the state θ. The prior of the state follows an improper uniform distribution over the real line.[9] There is a group of agents, indexed by $i \in \{1, 2 .., n\}$, who are privately and heterogeneously informed about θ. Agents' information set consists of two signals: (1) a common signal that is observed by all agents, which we represent by the random variable $y=\theta+\eta$ and (2) a private signal $x_i=\theta+\varepsilon_i$, which is only observed by agent i: We assume that the error terms $\{\{\varepsilon_i\}_{i=1}^n, \eta\}$ are mutually independent and jointly normally distributed with $\eta \sim N(0, \frac{1}{\alpha})$ and $\varepsilon_i \sim N(0, \frac{1}{\beta})$.

The market designer observes neither the common nor the private signals. To collect this information, the market designer asks each agent to make a sealed prediction about θ, so that agents simultaneously and independently predict the value of θ. In exchange, the market designer commits to reward agents according to a publicly announced reward rule that defines how each agent is rewarded based on the prediction made, the realization of θ, and the predictions of other agents. In particular, let a_i be agent i's action or prediction. The payoff of agent i who predicted a_i when the state of nature is θ and the opponents' predictions are $\mathbf{a}_{-i} \equiv \{a_1, \ldots, a_{i-1}, a_{i+1}, \ldots, a_n\}$ is given by

$$u_i(\theta, \mathbf{a}_{-i}, a_i) \equiv -\delta(a_i - \theta)^2 - (1-\delta)(a_i - \bar{a}_n)^2, \tag{2.1}$$

where $\bar{a}_n = \frac{a_1+a_2+\ldots+a_n}{n}$ is the average prediction among all agents, which we refer to as the *consensus forecast*. We assume that agents are risk neutral, maximizing the expected payoff.

This reward rule has two components:

1 The first component, the *accuracy* term, $(a_i-\theta)\text{-}^2$, is a standard quadratic loss in the distance between the underlying state and the action. This payoff component represents the cost of forecast error. Like in prediction markets, this component induces agents to anchor their predictions on the fundamentals θ.[10]

2 The second component, the *beauty contest* term, $-(a_i-\bar{a}_n)^2$, represents the cost of being away from the consensus forecast. As in idea markets, this term introduces a coordination motive or *beauty contest* incentive that induces an agent to second guess opponents' beliefs.[11]

In the sequel, we refer to δ (respectively, $1-\delta$) as the prediction market intensity (respectively, idea market intensity). We refer to the case $\delta=1$ as a *pure prediction market* and $\delta=0$ as a *pure idea market*.[12]

The reward rule (1) can also be interpreted probabilistically, in which case δ represents the probability the market designer observes the value of θ and rewards agents on that basis, and $1-\delta$ represents the probability that the designer does not observe θ and thus rewards agents on the basis of the consensus forecast.[13]

The structure of the game is common knowledge. When all agents use symmetric linear forecasting rules, as in all equilibria discussed below, the designer's best estimate of the location of θ is the consensus forecast $\bar{a}_n$.

3 Informativeness of idea markets

Our purpose is to characterize the determinants of the informational properties of the consensus forecast $\bar{a}_n$: In our setting with normally distributed posterior beliefs on θ, a natural measure of market *informativeness* is given by the precision of the posterior distribution of the state conditional on the consensus forecast, $\gamma = 1/\text{var}(\theta|\bar{a}_n)$. The value of γ is bounded above by $\alpha + n\beta$, which results in the most informative scenario where the prior uncertainty about θ is reduced both by knowledge about the common signal y and also by all agents' private signals $\mathbf{x} \equiv \{x_i\}_{i=1}^n$.

We proceed by first considering the two polar cases of a pure idea market $\delta=0$ and a pure prediction market $\delta=1$.

3.1 Pure idea market

Idea markets may suffer from multiplicity of equilibria arising from agents' incentives to coordinate their actions. Restricting attention to symmetric equilibria where strategies are linear in the common and private signals, this multiplicity of equilibria is only present when private signal errors ε_i are imperfectly correlated. For instance, when agents' private signals are perfect, they are indifferent between coordinating through the common signal (i.e., $a_i=y$) or through the private signals (i.e., $a_i=x_i=\theta$). In fact, they are indifferent between either of these pure alternatives and any possible combination of y and x_i. But, in the general case when private signals are imperfectly correlated, agents must disregard their private information in equilibrium.

Proposition 1 When $\delta=0$, in any symmetric equilibrium in linear strategies, agents put no weight on their private signal. Hence, $\gamma=1/\text{var}(\theta|\bar{a}_n) \leq 1/\text{var}(\theta|y)=\alpha$.

The proposition follows as a corollary to Proposition 4. A pure idea market is thus very uninformative because agents' actions only convey common information, which implies that all the information contained in private signals is lost. In a pure idea market, the coordination motive overwhelms agents' actions; an agent's prediction is chosen not on the basis of fundamentals, but only on the basis of the expected actions by the opponents.

3.2 Pure prediction market

In a pure prediction market there is a linear equilibrium in which agents submit their best linear predictor of θ.

Proposition 2 When $\delta=1$, there is a unique symmetric equilibrium in linear strategies in which $a_i = E(\theta \mid y, s_i) = \frac{\alpha y+\beta x_i}{\alpha+\beta}$ for all $i=1,\ldots, n$. In this equilibrium, informativeness is given by $\gamma(\delta=1) = \frac{n(\alpha+\beta)^2}{\alpha n+\beta}$.

Proof. The consensus forecast is given by $\bar{a}_n = \frac{\alpha y+\beta \bar{x}_n}{\alpha+\beta}$ where $\bar{x}_n = \frac{x_1+\ldots+x_n}{n}$, so that $\text{var}(\bar{a}_n|\theta)=(\frac{1}{\alpha+\beta})^2(\frac{\alpha^2}{\alpha}+\frac{\beta^2}{n^2}\frac{n}{\beta})=\frac{\alpha n+\beta}{(\alpha+\beta)^2 n}$. Given that the prior is improper, we have $\text{var}(\theta|\bar{a}_n)=\text{var}(\bar{a}_n|\theta)$ (see Hartigan, 1983) and thus we obtain the expression for the informativeness $\gamma(\delta=1)$.

In pure prediction markets, an agent's prediction is his individually most efficient estimate of the fundamentals θ. The consensus forecast is a weighted average of the common signal y and the sufficient statistic $\bar{x}_n$ for the private signals. As is well known, the consensus forecast puts greater weight on the common signal y than would an optimal estimate $E(\theta \mid y,\bar{x}_n) = \frac{\alpha y+n\beta x_n}{\alpha+n\beta}$ for θ based on y and $\bar{x}_n$.

The precision of this consensus estimate is always favored by the precision of both common and private information. Despite this, the informativeness of the consensus forecast may decrease in the precision of common information.

Corollary 3 When $\delta=1$, informativeness increases in β and in n but decreases in the precision of the common signal if and only if $(\alpha-\beta)n+2\beta\leq 0$.

Proof. From the expression in Proposition 2, we have $\frac{\partial\gamma(\delta=1)}{\partial\beta} = \frac{n(\alpha+\beta)[(2n-1)\alpha+\beta]}{(\alpha n+\beta)^2}$, $\frac{\partial\gamma(\delta=1)}{\partial\alpha} = \frac{n(\alpha+\beta)[n\alpha+(2-n)\beta]}{(\alpha n+\beta)^2}$, and $\frac{\partial\gamma(\delta=1)}{\partial n} = \frac{\beta(\alpha+\beta)^2}{(\alpha n+\beta)^2}$.

The excessive weight accorded by the consensus forecast to the common signal relative to the optimal predictor is particularly great when the number of agents is large. A higher α can aggravate this problem. If we held agents' strategies constant, an increase in α would always result in a more informative consensus forecast, simply because the consensus forecast would contain less noise. But an increase in α induces agents to assign additional weight to the common signal y, thereby reducing even more the correlation between the consensus forecast and agents' private signals. This effect would never reduce the informativeness of the consensus forecast when the common signal is more precise than the private signal, but it does so when the opposite is true.

4 Mixed incentives

The polar cases of pure markets discussed above serve as benchmarks. We now turn to the more interesting case in which agents have a mix of prediction market and idea market incentives, so that $\delta \in (0, 1)$. For convenience, we use the shorthand notation $E_i(\cdot) = E_i(\cdot|y, x_i)$.

Agent i's optimization program is to solve

$$\max_a \; E_i[-\delta(a-\theta)^2 - (1-\delta)(a-\bar{a}_n)^2], \tag{2.2}$$

and the agent's optimal action is characterized by the first-order condition

$$a_i = \delta E_i(\theta) + (1-\delta) E_i(\bar{a}_n). \tag{2.3}$$

Intuitively, equation (2.3) shows that agent i's optimal prediction is a weighted average of the assessment of the fundamentals and of the consensus forecast. To estimate the consensus forecast, an agent must forecast the forecasts of others, as in Townsend (1983).

To solve this problem, we first assume (and then verify) the existence of a symmetric linear equilibrium in which an agent's strategy is a weighted average of private and common signals, as described by

$$a_i = \phi y + (1-\phi) x_i \tag{2.4}$$

where φ is the weight attached to the common signal in an agent's equilibrium prediction. According to this candidate equilibrium, the agent expects the consensus forecast to be given by

$$E_i(\bar{a}_n) = \frac{a_i + (n-1)E_i(a_{-i})}{n} \tag{2.5}$$

where a_{-i} is the prediction of any opponent of agent i. Substituting $E_i(\bar{a}_n)$ into equation (2.3), we find

$$a_i = \frac{\delta E_i(\theta) + (1-\delta)\frac{(n-1)}{n} E_i(a_{-i})}{1 - \frac{(1-\delta)}{n}}$$

and, using (2.4), we obtain

$$a_i = \frac{\delta \frac{\alpha y + \beta x_i}{\alpha+\beta} + (1-\delta)\frac{(n-1)}{n}\left[\phi y + (1-\phi)\frac{\alpha y + \beta x_i}{\alpha+\beta}\right]}{1 - \frac{(1-\delta)}{n}}. \tag{2.6}$$

Finally, comparing coefficients in (2.4) and (2.6) yields

$$\phi = \frac{(n-1+\delta)\alpha}{(n-1+\delta)\alpha + n\delta\beta}, \tag{2.7}$$

thus establishing the following result:

Proposition 4 For all $\delta \in [0, 1]$, there exists a unique symmetric equilibrium in linear strategies $a_i = \varphi y + (1-\varphi)x_i$, where

$$\phi = \frac{(n-1+\delta)\alpha}{(n-1+\delta)\alpha + n\delta\beta}, \tag{2.8}$$

To introduce the properties of the equilibrium, we first consider two limit cases in terms of the number of participants in the market. When the number of agents grows large, the weight attached to the common signal in agents' equilibrium strategy converges to $\lim_{n\to\infty} \phi = \frac{\alpha}{\alpha+\delta\beta}$, which is larger than the weight prescribed by the individually optimal linear predictor, $\frac{\alpha}{\alpha+\beta}$. By contrast, when there is only one agent, the weight attached to the common signal coincides with that of the optimal linear prediction, $\lim_{n\to 1} \phi = \frac{\alpha}{\alpha+\beta}$. Intuitively, these cases suggest that agents tend to assign an excessive weight to the common signal when they have coordination concerns. The next corollary studies the determinants of ϕ.

Corollary 5 The weight ϕ attached to the common signal y in agents' equilibrium strategies decreases in the intensity of prediction markets δ and in the precision of the private signal β and increases in the number of agents n and in the precision of the common signal α.

Proof. We have $\frac{\partial\phi}{\partial\delta} = \frac{-n(n-1)\alpha\beta}{[(n-1+\delta)\alpha+n\delta\beta]^2} < 0$ for all $n > 1$, $\frac{\partial\phi}{\partial n} = \frac{\delta(1-\delta)\alpha\beta}{[(n-1+\delta)\alpha+n\delta\beta]^2} > 0$ for $\delta \in (0, 1)$, $\frac{\partial\phi}{\partial\alpha} = \frac{n(n-1+\delta)\delta\beta}{[(n-1+\delta)\alpha+n\delta\beta]^2} > 0$, and $\frac{\partial\phi}{\partial\beta} = \frac{-n(n-1+\delta)n\delta\alpha}{[(n-1+\delta)\alpha+n\delta\beta]^2} < 0$.

The effect of α and β on ϕ are intuitive: the weight attached to the common signal increases in the relative precision of the common signal.

The other two effects are more central to our questions regarding idea markets. Agents' coordination incentives become stronger when the intensity δ of prediction markets is lower, or when the size n of the market is larger, to the extent that agents are more concerned about coordination, so they increase the weight of the common signal. This weight is excessive relative to the optimal predictor of θ already in the pure prediction market, and the effect is only aggravated when more weight is accorded to the idea market.

4.1 Informativeness of consensus forecast

In this section we examine the determinants of informativeness. First, we consider how the precision of agents' signals affect informativeness, and then we study how the intensity of prediction markets and the number of agents affect informativeness. We show that both the precision of common information and the number of agents may be detrimental to informativeness. As expected, we verify that the intensity of prediction markets enhances the informativeness of the consensus forecast.

Our analysis relies on the following intermediate result:

Lemma 6 The posterior belief $\theta|\bar{a}_n \sim N(\bar{a}_n, \mathrm{var}(\theta|\bar{a}_n))$ where

$$\mathrm{var}(\theta\,|\,\bar{a}_n) = \frac{\phi^2}{\alpha} + \frac{(1-\phi)^2}{n\beta} = \frac{(n-1+\delta)^2\alpha + n\delta^2\beta}{[(n-1+\delta)\alpha + n\delta\beta]^2}. \qquad (2.9)$$

Proof. Since $\bar{a}_n = \frac{a_1+\ldots+a_n}{n} = \phi y + (1-\phi)\frac{x_1+\ldots+x_n}{n}$, we have $\bar{a}_n|\theta \sim N(\theta, \mathrm{var}(\bar{a}_n|\theta))$ with $\mathrm{var}(\bar{a}_n|\theta) = \phi^2\frac{1}{\alpha} + (1-\phi)^2\frac{n}{n^2\beta}$. It follows from Hartigan (1983) that the posterior is as claimed with $\mathrm{var}(\theta|\bar{a}_n) = \mathrm{var}(\bar{a}_n|\theta)$. The second part of equation (2.9) follows from (2.7).

Armed with this lemma, we now turn to the determinants of informativeness. We begin by investigating the impact of the most direct determinants of informativeness, the precision of the signals.

Proposition 7 (1) Informativeness increases with the precision of the private signal β; (2) Informativeness increases with the precision of the common signal α if and only if $\alpha/\beta \geq \frac{(n-1-\delta)n\delta}{(n-1+\delta)^2}$.

Proof. For part (1) we use (2.9) to find $\frac{\partial\,\mathrm{var}(\theta|\bar{a}_n)}{\partial\beta} = -n\delta\frac{[\delta(\delta+3n-3)+2(n-1)^2]\alpha + n\delta^2\beta}{[(n-1+\delta)\alpha+n\delta\beta]^3} < 0$.

For part (2), we can likewise derive $\frac{\partial\,\mathrm{var}(\theta|\bar{a}_n)}{\partial\alpha} = \frac{(n-1+\delta)[(n-1-\delta)n\delta\beta-(n-1+\delta)^2\alpha]}{[(n-1+\delta)\alpha+n\delta\beta]^3}$. Its sign depends on the factor $(n-1-\delta)n\delta\beta-(n-1+\delta)^2\alpha$, as stated.

Corollary 3 already established that even in a pure prediction market, a more precise common signal may reduce informativeness, simply because the consensus forecast may end up capturing too little of the information contained in the private signals. The presence of idea market incentives should, if anything, aggravate this problem because agents' coordination incentives induce them to overweight the common signal, thereby increasing the redundancy of common information in the consensus forecast.

Consider the effect of the intensity of idea markets on informativeness.

Proposition 8 Informativeness decreases in the intensity of idea markets: $\frac{\partial\,\mathrm{var}(\theta|\bar{a}_n)}{\partial\delta} > 0$

Proof. Note that $\text{var}(\theta\,|\,\bar{a}_n) = \frac{\phi^2}{\alpha} + \frac{(1-\phi)^2}{n\beta}$ is a convex function of φ which has a unique minimum at $\underline{\phi} = \frac{\alpha}{\alpha+n\beta}$, and therefore is increasing for $\varphi > \underline{\varphi}$. By (2.7), $\phi > \underline{\phi}$ is equivalent to $(n-1+\delta) > \delta$, which is true. From Corollary 5, $\frac{\partial\phi}{\partial\delta} < 0$, we can conclude that $\text{var}(\theta|\bar{a}_n)$ increases with δ.

The reason the intensity of idea markets hampers the informational content of the consensus forecast should not be surprising. The presence of idea market incentives excessively reduces the correlation between agents' predictions and agents' information about fundamentals. A clear implication from this observation is that a market designer with discretion over the value of δ might want to set $\delta = 1$ in order to maximize the informational content of the consensus forecast. We conclude that a designer who observes *ex post* the value of the fundamentals would optimally organize this market as a pure prediction market.

The informational effect of n is less intuitive and perhaps more surprising. Increasing the number of agents has two effects. On the one hand, there are the traditional statistical benefits of a larger sample: holding agents' strategies constant, an increase in the number of agents allows the consensus forecasts to filter more efficiently the noise included in agents' private signals. We call this the *large sample effect*. On the other hand, as the number of agents increases, agents are induced to strategically increase the weight they assign to the common signal, thereby reducing the informational content of the consensus forecast. This is the *second guessing effect*. Remarkably, under certain circumstances the second guessing effect overwhelms the large sample effect so that increasing n decreases the informational content of $\bar{a}_n$.

Proposition 9 For sufficiently small prediction market intensity, $\delta < \delta^* = \frac{\sqrt{9\alpha^2+8\alpha\beta}-3\alpha}{2\beta} \in (0,\,1)$, informativeness is a single-peaked function of the number of agents n, decreasing when $n > n^* = \frac{(1-\delta)(2-\delta)\alpha}{(2-3\delta)\alpha-\delta^2\beta}$. When instead $\delta \geq \delta^*$, informativeness increases in n.

Proof. From (2.9), we have $\frac{\partial\,\text{var}(\theta|\bar{a}_n)}{\partial n} = \beta\delta\frac{(\delta-1)(2-\delta)\alpha+[(2-3\delta)\alpha-\delta^2\beta]n}{[(n-1+\delta)\alpha+n\delta\beta]^3}$. Its sign depends on $A=(\delta-1)(2-\delta)\alpha+[(2-3\delta)\alpha-\delta^2\beta]n$, which has a negative intercept, $(\delta-1)(2-\delta)\alpha<0$. The slope $\frac{\partial A}{\partial n}=(2-3\delta)\alpha-\delta^2\beta$ is a concave function of δ and takes the value of $2\alpha>0$ when $\delta=0$ and the value $-(\alpha+\beta)<0$ when $\delta=1$. Given that $\frac{\partial A}{\partial n}$ attains its maximum at $\delta=-\frac{3\alpha}{2\beta}<0$ and decreases in δ over $[0,\,1]$, there exists $\delta^*\in(0,\,1)$ such that $\frac{\partial A}{\partial n}\leq 0$ if and only if $\delta\geq\delta^*$, as stated. Combining this fact with the earlier observation that the intercept is negative, we conclude that $A<0$ if $\delta\geq\delta^*$. By contrast, $\frac{\partial A}{\partial n}>0$ if $\delta<\delta^*$, in which case $A\geq 0$ if and only if $n\geq n^*$, as stated.

Figure 2.1 plots informativeness γ as a function of the number of agents n. There we see that a larger sample of agents increases the information content of the consensus forecast when the intensity of idea markets is low, as illustrated by the solid curve. By contrast, a larger sample aggravates the second guessing effect when the intensity of idea markets is high, as in the dashed and dotted curves. This is remarkable, because it means that the second guessing effect associated with the intensity of idea markets may destroy a lot of information, even reversing the benefits of a larger sample of agents.

5 Conclusion

In this chapter we characterize the informational properties of the consensus forecast when agents trade off the incentive to accurately predict fundamentals (as in pure prediction markets) with the desire of meeting the most popular opinion (as in pure idea markets, or beauty contests).

The main lesson of our analysis is that many of the intuitive properties that usually characterize standard markets or even prediction markets do not hold when agents have coordination motives, as in idea markets. For example, neither a larger sample of agents nor a better quality of common information are necessarily beneficial in terms of the informational properties of the market. The design of an idea market is thus a delicate task.

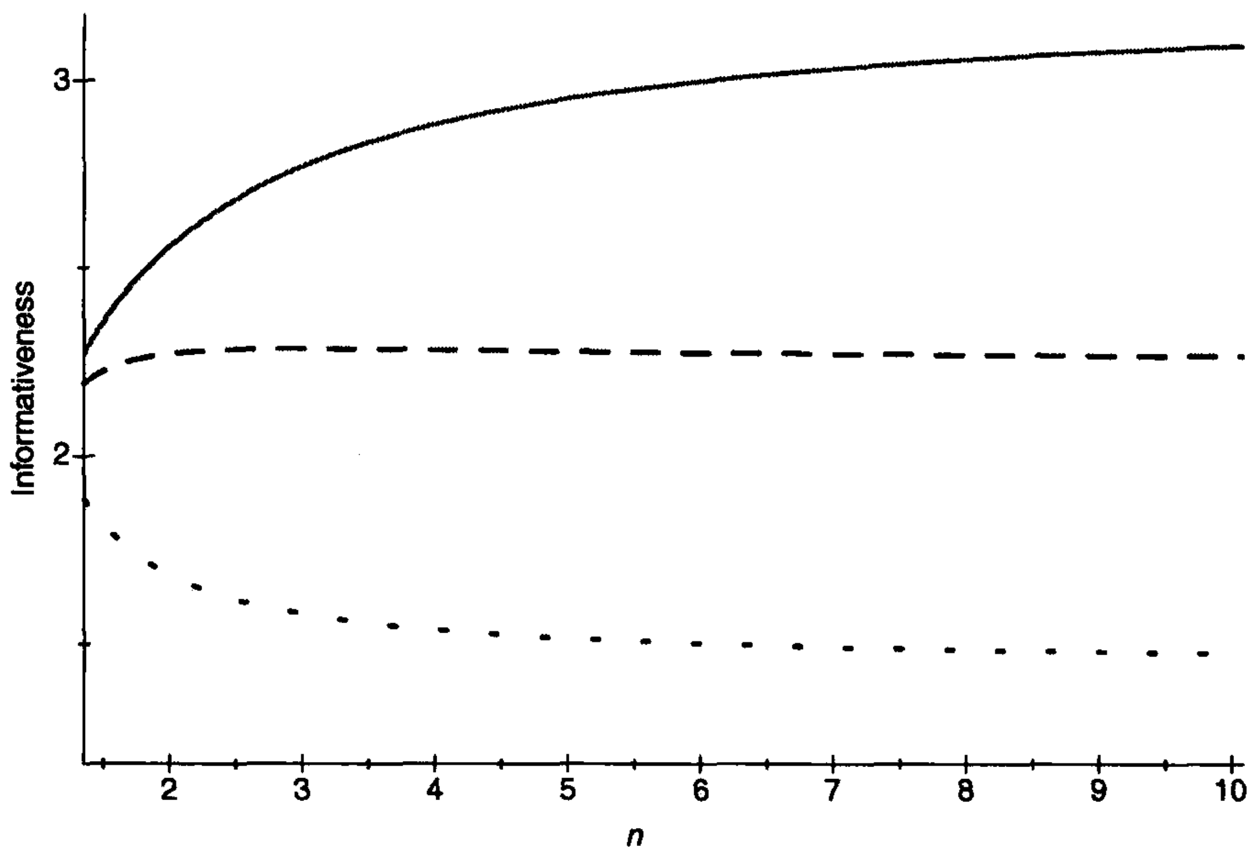


Figure 2.1 Illustration of the impact of the number of agents n on informativeness γ for $\alpha=\beta=1$. The curves correspond to $\delta=0.8>\delta^*$ (solid), $\delta=0.5<\delta^*$ (dashed, with corresponding $n^*=3$), and $\delta=0.1$ (dotted, with corresponding $n^*\approx 1$).

A key limitation of our analysis is that it relies on reduced-form rewards, in keeping with the literature on beauty contests. A crucial assumption in our model is that agents' actions are strategic complements. It would be interesting to analyze idea markets in a fully micro-founded model of trading. We expect our results to be affected if we instead assume that agents' actions are strategic substitutes (see Angeletos and Pavan, 2007), as in financial markets where traders have market power à la Kyle (1985).

Notes

1 See Plott and Chen's (2002) pioneering experiment. For a brief introduction to the topic we refer to Ottaviani (2009).
2 See Wolfers and Zitzewitz (2004) for an introduction to prediction markets. See Berg and Rietz (2006) and Forsythe *et al.* (1992) on the Iowa electronic markets. See Ottaviani and Sørensen (2009, 2010) and references therein on horse-race betting.
3 See Hayek's (1945) classic essay about the ability of markets to aggregate dispersed information, Grossman's (1976) theoretical development of the concept of rational expectation equilibrium, and Surowiecki's (2004) popularization of how markets can harness the wisdom of the crowds.
4 See Spears *et al.* (2009).
5 Wolfers and Zitzewitz (2006) also stress that the key variables of interest are often difficult to write into contracts.
6 See Keynes (1936, p. 156):

> professional investment may be likened to those newspaper competitions in which the competitors have to pick out the six prettiest faces from a hundred photographs, the prize being awarded to the competitor whose choice most nearly corresponds to the average preferences of the competitors as a whole; so that each competitor has to pick, not those faces which he himself finds prettiest, but those which he thinks likeliest to catch the fancy of the other competitors, all of whom are looking at the problem from the same point of view. It is not a case of choosing those which, to the best of one's judgement, are really the prettiest, nor even those which average opinion genuinely thinks the prettiest. We have reached the third degree where we devote our intelligences to anticipating what average opinion expects the average opinion to be. And there are some, I believe, who practise the fourth, fifth and higher degrees.

7 In a pure idea market, we find that there is a unique equilibrium in linear strategies, contrary to conventional wisdom. Multiplicity arises only when agents have common knowledge about fundamentals. But the slightest breakdown in agents' common knowledge results in a unique equilibrium, suggesting that multiplicity of equilibria is a knife-edge result.
8 This is a feature also stressed by Morris and Shin (2002). However, by focusing on a version of the model with a continuum of agents, Morris and Shin do not analyze how information aggregation depends on the prediction market intensity and the number of informed agents. Instead, we focus on these questions that are natural in the context of our interpretation in terms of idea markets.
9 The assumption that θ is uniformly distributed on the real line is non-standard, but presents no technical difficulties as long as we are concerned with conditional beliefs. Morris and Shin (2002) argue that this assumption can be considered as the limiting case as the prior distribution of θ becomes diffuse.
10 This is a reduced-form model of a market, but we may imagine agent i adopting a position in a risky asset. The ideal position depends on the parameter θ (or $\bar{a}_n$ if the

designer substitutes that for θ in order to close market positions). By optimality, there is a negligible first-order loss of trading as if the parameter is equal to a_i and our payoff expression $-(a_i-\theta)^2$ captures the second-order effect.

11 The fact that agents have a coordination motive means that their actions are strategic complements. Note, however, that trading in a financial market might induce strategic substitutability. Also, the incentive to be close to the consensus is opposite to the incentive to differentiate one's forecast in a large winner-takes-all forecasting contest (see Ottaviani and Sørensen, 2006). To properly illustrate the consequences of a dominant beauty contest effect, here we focus on the case where the total effect results in strategic complements.

12 Of course, this simple reward rule is somewhat arbitrary, and may not even be optimal, but provides a simple setting to compare the informational properties of prediction and idea markets.

13 In general, one would think that δ is endogenous as it depends on whether or not the market designer undertakes the idea. This is perhaps an interesting extension.

References

Angeletos, George-Marios and Alessandro Pavan (2007), "Efficient Use of Information and Social Value of Information," *Econometrica*, 75(4), 1103–1142.

Berg, Joyce E. and Thomas A. Rietz (2006), "The Iowa Electronic Markets: Stylized Facts and Open Issues." In *Information Markets: A New Way of Making Decisions*, edited by Robert W. Hahn and Paul C. Tetlock. AEI-Brookings Joint Center for Regulatory Studies.

Forsythe, Robert, Forrest Nelson, George R. Neumann and Jack Wright (1992), "Anatomy of an Experimental Political Stock Market," *American Economic Review*, 82(5), 1142–1161.

Grossman, Sanford J. (1976), "On the Efficiency of Competitive Stock Markets where Traders have Diverse Information," *Journal of Finance*, 31(2), 573–585.

Hahn, Robert W. and Paul C. Tetlock (2007), "Optimal Liquidity Provision for Decision Makers." Working Paper, University of Texas at Austin and AEI-Brookings Joint Center for Regulatory Studies.

Hanson, Robin and Ryan Oprea (2009), "A Manipulator can Aid Prediction Market Accuracy," *Economica*, 76(302), 304–314.

Hartigan, John (1983), *Bayes Theory*. Springer, New York.

Hayek, Friedrich A. (1945), "The Use of Knowledge in Society," *American Economic Review*, 35(4), 519–530.

Keynes, John Maynard (1986), *The General Theory of Employment, Interest and Money*, Macmillan.

Kyle, Albert S. (1985), "Continuous Auctions and Insider Trading," Econometrica, 53(6), 1315–1335.

Lieli, Robert P. and Augusto Nieto-Barthaburu (2009), "Feedback in Prediction Markets," Working Paper, University of Texas at Austin and East Carolina University.

Morris, Stephen and Hyhn Song Shin (2002), "Social Value of Public Information," *American Economic Review*, 92(5), 1521–1534.

Ottaviani, Marco (2009), "The Design of Idea Markets: An Economist's Perspective," *Journal of Prediction Markets*, 3(1), 41–43.

Ottaviani, Marco and Peter Norman Sørensen (2006), "The Strategy of Professional Forecasting," *Journal of Financial Economics*, 81(2), 441–466.

Ottaviani, Marco and Peter Norman Sørensen (2007), "Outcome Manipulation in Corporate

Prediction Markets," *Journal of the European Economic Association, Papers and Proceedings*, 5(2–3), 554–563.

Ottaviani, Marco and Peter Norman Sørensen (2009), "Surprised by the Parimutuel Odds?" *American Economic Review*, 99(5), 2129–2134.

Ottaviani, Marco, and Peter Norman Sørensen (2010), "Noise, Information, and the Favorite-Longshot Bias in Parimutuel Predictions," *American Economic Journal: Microeconomics*, 2(1), 58–85.

Plott, Charles and Kay-Yut Chen (2002), "Information Aggregation Mechanisms: Concepts, Design and Implementation for a Sales Forecasting Problem," Social Science Working Paper 1131, California Institute of Technology.

Spears, Brian, Christina LaComb, John Interrante, Janet Barnett and Deniz Senturk-Dogonaksoy (2009), "Examining Trader Behavior in Idea Markets: An Implementation of GE's Imagination Markets," *Journal of Prediction Markets*, 3(1), 17–39.

Surowiecki, James (2004), *The Wisdom of Crowds*. Doubleday Publishing, New York.

Townsend, Robert M. (1983), "Forecasting the Forecasts of Others," *Journal of Political Economy*, 91(4), 546–588.

Wolfers, Justin and Eric Zitzewitz (2004), "Prediction Markets," *Journal of Economic Perspectives*, 18(2), 107–126.

Wolfers, Justin and Eric Zitzewitz (2006), "Five Open Questions about Prediction Markets." In *Information Markets: A New Way of Making Decisions*, edited by Robert W. Hahn and Paul C. Tetlock. AEI-Brookings Joint Center for Regulatory Studies.

3 How prediction markets can save event studies

Erik Snowberg, Justin Wolfers and Eric Zitzewitz

1 Introduction

Event studies have been used in political science to study the cost of regulation (Schwert, 1981), the value of political connections (Roberts, 1990a; Fisman, 2001), the effect of political parties on defense spending (Roberts, 1990b), the importance of rules in congressional committees (Gilligan and Krehbiel, 1988), the reaction of different interests to trade legislation (Schnietz, 2000), how party control in parliamentary systems affects broad-based stock indices (Herron, 2000), the value of defense contracts (Rogerson, 1989), the effect of the political party of the US President and congressional majorities on particular industry segments (Mattozzi, 2008; Knight, 2006; Herron *et al.*, 1999; Den Hartog and Monroe, 2008; Monroe, 2008; Jayachandran, 2006), and other questions.

Yet event studies are under-used and on the decline. Many of the studies cited above mention the paucity of event studies in political science, despite their power in revealing the economic impacts of political events. Indeed, due to the issues detailed in this chapter, event studies have even fallen out of favor in their traditional strongholds: economics and finance. In what follows we demonstrate, in political contexts, several problems with event studies. We then show that by pairing traditional event studies with prediction markets, to form a *prediction market event study*, many of the problems that have plagued event studies can be mitigated, clearing the way for broader application of event studies in political science.

At its core, an event study tries to ascertain the economic impacts of a particular event, such as a policy change. These impacts are then used to judge political theories that make different predictions about either the direction or size of an economic effect of a political event, or how an event will differentially affect different economic sectors.

To begin an event study, a researcher first chooses which economic indicator(s) he or she is interested in. If the researcher is interested in the effect of a policy change on a company's performance, he or she will likely examine the price of that company's stock s.[1] The researcher will then select some time t before the policy change occurred, and some time t' after the policy change. The period between t and t' is called the *event window*.

The difference in the price of the stock between the beginning and the end of the event window is proportionate to the impact of the policy change on the price of the company's stock. However, policy changes rarely appear suddenly. Instead, they are created over many months or years. In order to judge the full impact of the policy change on the company's stock, the researcher must judge the *prior probability* of the event at the beginning of the window. We call the chosen prior probability π. Since the posterior probability of the event is always 1, the full effect E of the policy change on the company's stock is given by:[2]

$$E = \frac{s_{t'} - s_t}{1 - \pi}, \tag{3.1}$$

where s_t is the price of the company's stock at time t.

It should be clear that if the researcher chooses a different start time t, end time t', or prior probability π, the event study can produce vastly different results in terms of the size and even sign of the estimated effect E. In particular, the longer the event window, the more likely it is that other, unrelated events will occur which will bias the results.

Prediction markets can mitigate these problems. While prediction markets may take many forms, here we focus on the market for a contract that pays \$1 if a certain event, such as a policy change, happens and zero otherwise.[3] The price at any given time thus represents the market estimate of the probability of that event happening.[4]

Prediction markets give the researcher an accurate measure of π, the prior probability of an event happening, and may also help to identify an appropriate event window. Moreover, changes in the company's stock price that are unrelated to changes in the probability of the policy change will show no correlation with changes in the price of the prediction market contract. Thus, the event window does not need to be carefully chosen to exclude other events. Finally, as the probability of an event may change many times in response to political events, each probability change can be analyzed as a separate event, where the change in probability of the policy change is accurately measured. By separating a single event window into many small sub-windows and then taking an appropriately weighted average effect across the sub-windows, prediction markets allow for more precise estimates than traditional event studies.

This chapter consists of three examples from our previous research (Snowberg *et al.*, 2007a,b).[5] Each example illustrates a particular problem with traditional event studies and shows how the inclusion of prediction markets produces a more accurate estimate of the economic effect of political events. It should be noted that the three issues with traditional event studies are all inter-related – so each example will contain some elements of the other issues. Where possible, we show the difference between our results and research using traditional event studies or misusing prediction markets.

The examples in this chapter provide insight into several questions, as well as demonstrating the methodological usefulness of prediction markets. First, we show that in the 2004 US presidential election, candidate convergence did not occur, as predicted by Downs (1957) and many other models. Specifically, the stock market rose 2 percent in value on news of a Bush victory (over Kerry). Second, we show this difference of 2 percent between Republicans and Democrats has been remarkably consistent over time, appearing in an analysis of all elections between 1880 and 2004. This suggests that whatever the changes in party structure and policy issues over that period, Republicans have consistently been the party of capital, and Democrats the party of labor. Finally, we show that the stock market declined in response to the news of a Democrat victory in the Senate (and House) in 2006, suggesting that, contrary to conventional wisdom, markets do not prefer divided control of the legislature and executive to unified control of both branches.

2 Choosing the event window

To illustrate the importance of the event window in traditional event studies, we focus on election night 2004. The price of an S&P 500 future, and the InTrade prediction market tracking Bush's probability of re-election are shown in Figure 3.1.[6] Around 3 p.m. EST on election day, the S&P 500 and the probability of Bush's re-election declined in response to leaked exit polls. These polls, which were improperly analyzed (or not analyzed at all) by the news media, showed a Kerry lead in many battleground states. As the actual election results were tallied and Bush's victory became apparent, both the S&P 500 and the probability of Bush's re-election (as revealed by the prediction market) rose.

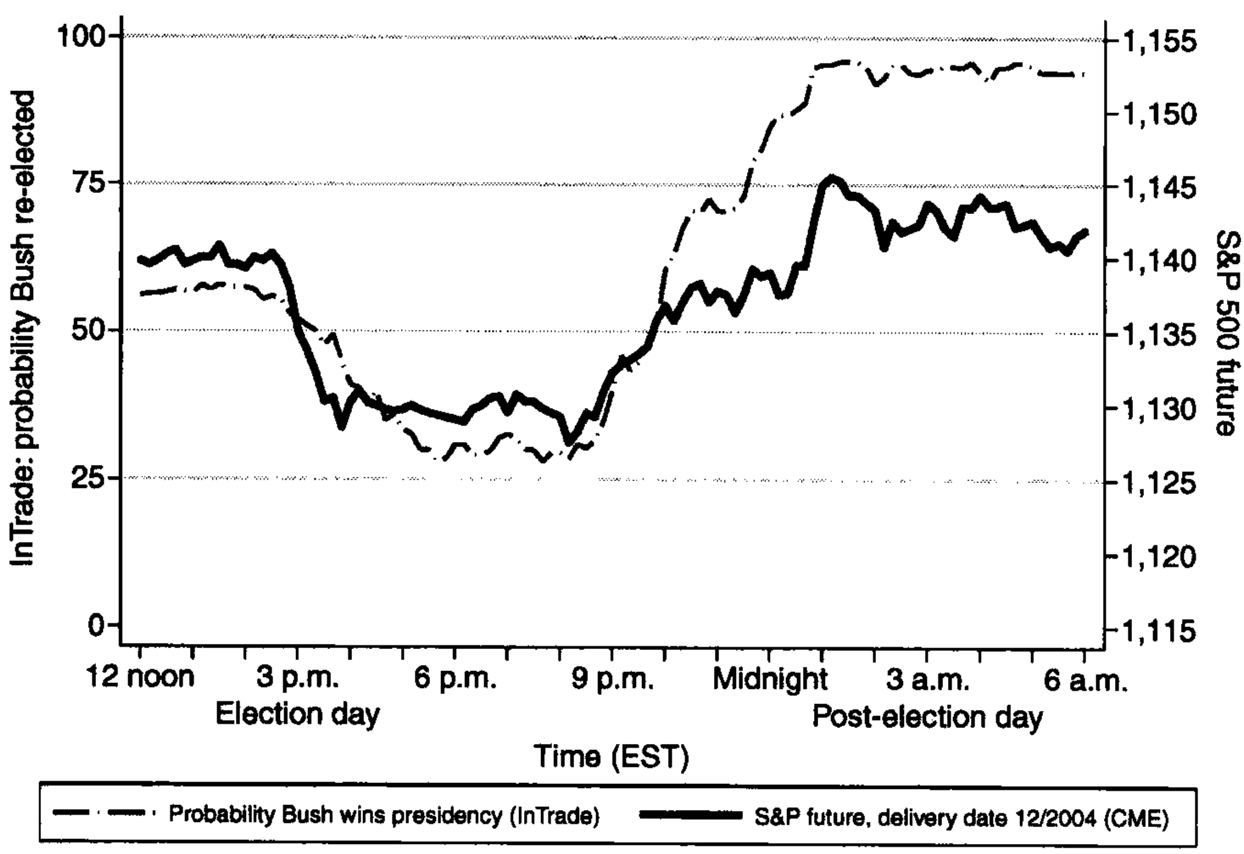


Figure 3.1 Prediction markets reveal that Bush's re-election on November 2, 2004 increased equity prices by approximately 2%.

Combining data from both the prediction market and the S&P 500 makes it clear that the S&P 500 rose in response to a Bush victory over Kerry. Moreover, the prediction market allows us to fix an exact scaling between the rise in the probability of a Bush presidency and the rise in the S&P 500. For example, in the first natural experiment in Figure 3.1 from 3–5 p.m., the probability of a Bush victory decreased by approximately 30 percent, while the S&P 500 decreased by approximately 0.7 percent. Plugging these values into (3.1) (appropriately modified so that the denominator reflects the probabilities before and after the event, rather than a value of 1 after the event) yields that the S&P 500 would have declined by 2.3 percent on news of a Kerry victory (over Bush). Repeating the same exercise for the second natural experiment between 8 p.m. and 1 a.m. the next day, when Bush's chances of re-election rose 65 percent and the S&P 500 rose 1.3 percent, implies that the market rose 2.0 percent on news of a Bush victory. Repeating the same exercise every half hour for the time period in Figure 3.1 and averaging suggests that the market rose 2.1 percent on news of a Bush victory (standard error 0.5 percent).[7]

Without the information from prediction markets, what might a researcher make of the S&P 500's movement on election night? As previously mentioned, a researcher would have to choose the beginning and end of an event window, as well as the probability that Bush had of winning the presidency at the beginning of the event window. While it is hard to predict what any given researcher would choose to do in this situation, there are some points that are focal. For example, the opening of the market on election day at 9:30 a.m., the point immediately before the release of any exit polling at 3 p.m. and the close of polls in many east coast states at 8 p.m. are focal.

What about the end of the event window? The most obvious end point is when the major networks called the election for Bush. However, the networks hesitated to call the election in 2004, so a better selection would be when the networks called Ohio for Bush, at 1 a.m. the day after the election.

Two other points are focal end points: the opening of the market the day after the election (9:30 a.m.) and the closing of the market the day after the election (4:00 p.m. – which would give the markets more time to fully respond to the election news).[8] We will examine the effect of the prior probability of Bush's victory in the next section: for now we set it at 50 percent, as the election was generally considered to be closely contested up until election day.

Using (3.1), the middle column of Table 3.1 presents the effect of electing Bush, rather than Kerry, on the S&P 500 for the nine different possible event windows defined by focal start and end points. These nine different event windows provide nine different estimates that range from –1.5 percent to 2.6 percent. The standard errors are also much larger than those of the prediction market event study. The highest t-statistic in the table is 1.5, which implies that that estimate is not statistically significant at the 10 percent level. That is, none of the estimates in Table 3.1 are significant at conventional levels.

One event window yields the same estimate as the prediction market method (2.1 percent). While this is heartening, note that this window is just one window

Table 3.1 Different event windows and prior probabilities of Bush's re-election give widely different estimates of the market's response to Bush's victory over Kerry

Event window		*Prior probability*		
Start (11/2)	*End (11/3)*	*30%*	*50%*	*90%*
	1 a.m.	0.6% (0.9%)	0.8% (1.3%)	3.9% (6.3%)
9:30 a.m.	9:30 a.m.	0.9% (1.0%)	1.2% (1.4%)	6.1% (6.9%)
	4 p.m.	−1.1% (0.8%)	−1.5% (1.2%)	−7.5% (5.9%)
	1 a.m.	1.0% (0.7%)	1.4% (1.0%)	7.0% (4.8%)
2 p.m.	9:30 a.m.	1.3% (0.9%)	1.8% (1.2%)	9.2% (6.0%)
	4 p.m.	−0.6% (0.5%)	−0.9% (0.7%)	−4.4% (3.6%)
	1 a.m.	1.5% (1.2%)	2.1% (1.6%)	10.6% (8.2%)
8 p.m.	9:30 a.m.	1.8% (1.2%)	2.6% (1.7%)	12.8% (8.5%)
	4 p.m.	−0.1% (1.0%)	−0.2% (1.4%)	−0.8% (6.8%)

Notes
Standard errors (in parenthesis) are the standard deviation of price movements in the near quarter CME S&P 500 future for 60 days before and after November 2, 2004, appropriately adjusted by the prior probability.

out of nine that could be focal, and the fact that this window is correct is somewhat random. Moreover, this is the second highest of the nine possible estimates (in the middle column). Further, three of the estimates in this column are negative – that is, our hypothetical researcher might even get the direction of the effect wrong; that is, that markets *rose* on news of a Kerry (rather than Bush) presidency. Finally, note that the availability of intraday data allows much greater flexibility in the choosing of event windows. This adds researcher discretion and potential bias in traditional event studies, while making prediction market event studies more precise.

3 Prior probabilities

In the previous section we set the prior probability of Bush's victory at 50 percent, assuming that our hypothetical researcher thought the election was a toss-up. Other hypothetical researchers might believe that Bush's probability of winning re-election was closer to 30 percent if they were a Kerry partisan or as high as 90 percent if they used the results of Fair (1978, 1996, 2004) as a prediction.

The first and third columns of Table 3.1 examine the same nine event windows as before for the prior probabilities of 30 percent *and* 90 percent. Of the 27 different possible answers from a traditional event study, only one gives the correct number: 2.1 percent. The range of possible answers is also quite wide, as it may be as low as −7.5 percent or as high as 12.8 percent. This 20 percentage-point range is roughly centered on the correct answer, but that doesn't give much hope that a traditional event study will give the correct number.

Event study estimates are significantly more noisy without prediction markets to help time and calibrate them. One might wonder, though, whether this noise is sufficiently mitigated by the law of large numbers when one has a large number of similar events and is only interested in the average event effect. For instance, if one looked at a large number of presidential elections and assumed a pre-election probability of 50 percent in each case, would these errors cancel out?

This is the approach of Santa-Clara and Valkanov (2003), which finds no difference between the stock market's reaction to a Democrat or Republican winning the presidency. Their methodology is equivalent to conducting an event study for each election between 1928 and 1996, defining the event window as the period between the market close the day before the election and the market close the day after the election.[9] They chose 50 percent as the prior probability that the Republican candidate will win each election, and use either 0 or 1 as the posterior probability of the Republican winning (as at the end of the event window the Republican will have either won or lost). They then average the results of these event studies to get a mean and standard deviation, which can be used for statistical tests.

This methodology is also equivalent to regressing the percentage change in the stock market during the event window on the change in the probability of a Republican victory during the event window. This is illustrated in the first panel of Figure 3.2. This results in a coefficient of 1.3 percent and a standard error of 0.9 percent. Thus, the coefficient is not statistically significant at conventional levels.

In the second panel of Figure 3.2, we instead use prior probabilities gleaned from the historical prediction markets described in Rhode and Strumpf (2004, 2006, 2008). These markets, run on the curb exchange on Wall Street, were the dominant form of election projection before the advent of scientific polling. We use these markets to establish the true prior probability of a Republican winning the presidency each election. In this panel, the relationship between the electoral fortunes of Republicans and the stock market is clearly much stronger. By using prediction market event studies, we find that the markets rose 3.0 percent in response to a Republican victory between 1928 and 1996. The standard error of this estimate is 1.2 percent, making the coefficient statistically significant at the 5 percent level.

Why are our results so different from those of Santa-Clara and Valkanov (2003)? Using a 50 percent prior probability of Republican victory for each election, even if correct on average, adds measurement error to the right-hand side of the regressions. This biases coefficients towards 0. By using an accurate estimate of the prior probabilities of Republican victory gleaned from prediction markets, we are able to recover the correct coefficient, which is similar to the result from 2004.

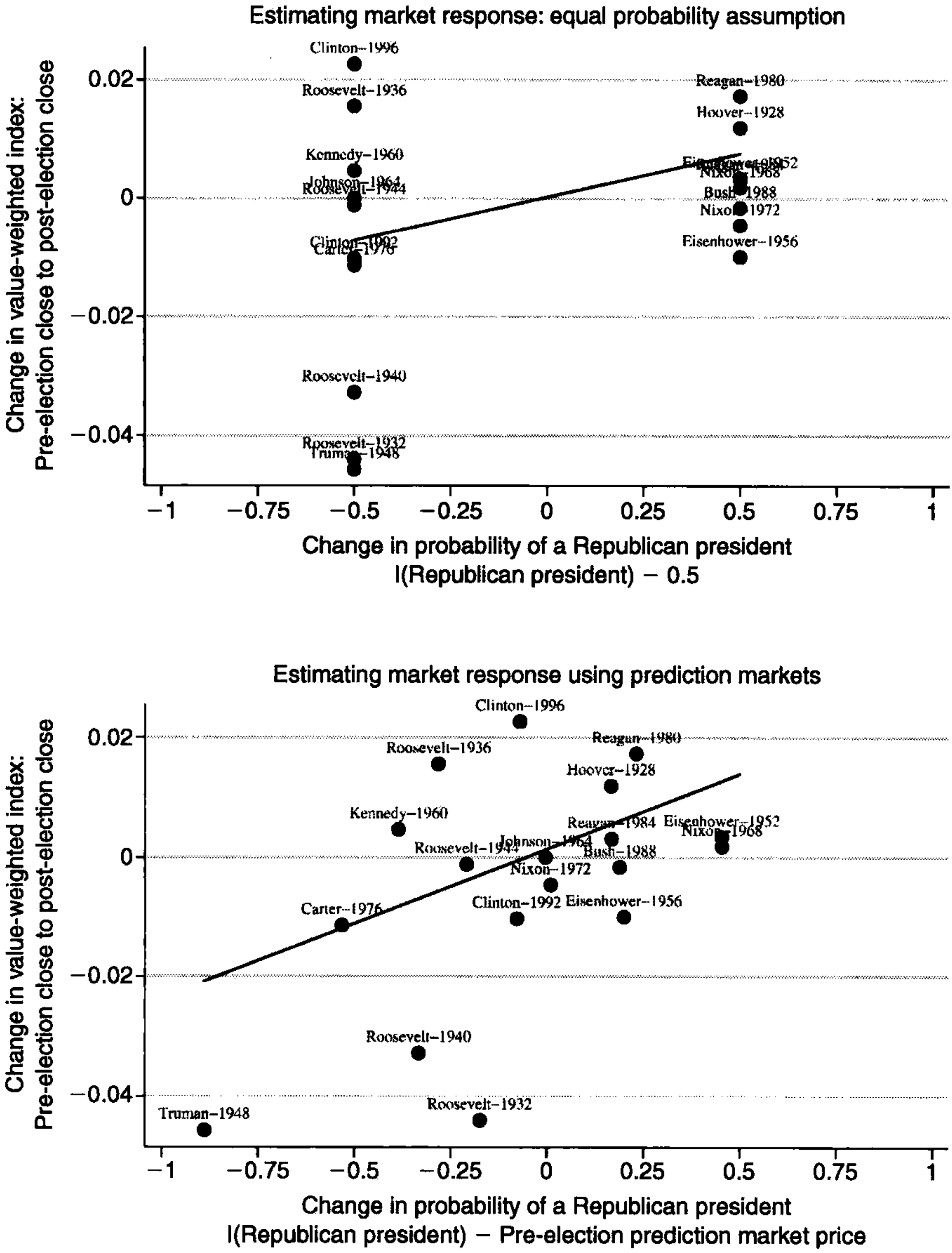


Figure 3.2 The long-term relationship between elections and equities is obscured without prediction markets.

4 Parceling out market movements related to politics

Even if a researcher picked the correct event window and prior probability, the results of a traditional event study could still be biased by market movements caused by other events. While financial theory says that, on average, market movements due to other events should have very little effect, the probability that they actually have no effect over a moderate length of time (a day) is quite small.

To illustrate this, consider the 2006 US congressional election, as illustrated in Figure 3.3. The first panel displays the entire course of election night, showing the movements of a prediction market contract tied to whether Republicans maintained a majority in the Senate, another contract tied to whether the Republicans maintained a majority in the House, and an S&P 500 near-month future.

At the beginning of election night, Republicans had less than a 20 percent chance of maintaining their majority in the House. This slowly converged to zero as the results of the few close contests became known. The Republicans in the Senate seemed to be in a better position to hold on to their majority, entering the evening with a 70 percent chance that rose above 90 percent when early exit polls favored them. However, when the vote totals of the last counties came in for Virginia and Missouri, the markets quickly reversed themselves, sending the probability of a Republican Senate majority down to 10 percent.

A cursory glance at these three financial contracts together would suggest that there was little relationship between Republican (vs. Democrat) majorities in the House and Senate and the S&P 500. The markets did trend downwards over the entire evening, which may have led a traditional event study (even with the proper prior probabilities) to conclude that the change of the House from a Republican to Democrat majority would result in a *decrease* in the S&P 500 of over 2 percent. However, if the event window included Rumsfeld's resignation, a traditional event study would have concluded that the change of the House from a Republican to a Democrat majority resulted in more than a 2 percent *increase* in the S&P 500.[10]

The truth, as revealed by prediction markets, is strikingly different. The second panel of Figure 3.3 shows a close-up of the time period when there was a large change in the probability of a Republican majority in the Senate. This panel shows that the market moves in lock-step with the probability of a Republican majority in the Senate, although the movement is small. Moreover, the market's non-response to the swing of Virginia's Senate seat into the Democrat column reveals that the movement in the S&P 500 is due to a change in the majority status of the parties, not a shift in the median voter of the Senate.

By employing the same methodology as that used in our study of the 2004 election, we find that a shift in the Senate from a Republican to a Democrat majority caused the S&P 500 to decline by 0.2 percentage points (standard error 0.09 percent – making the result statistically significant at the 10 percent level), while the S&P 500 was statistically indifferent to the change in the House majority.[11] It should be noted that these results are from a regression that includes the *entire* time covered by the first panel of Figure 3.3, further underscoring the

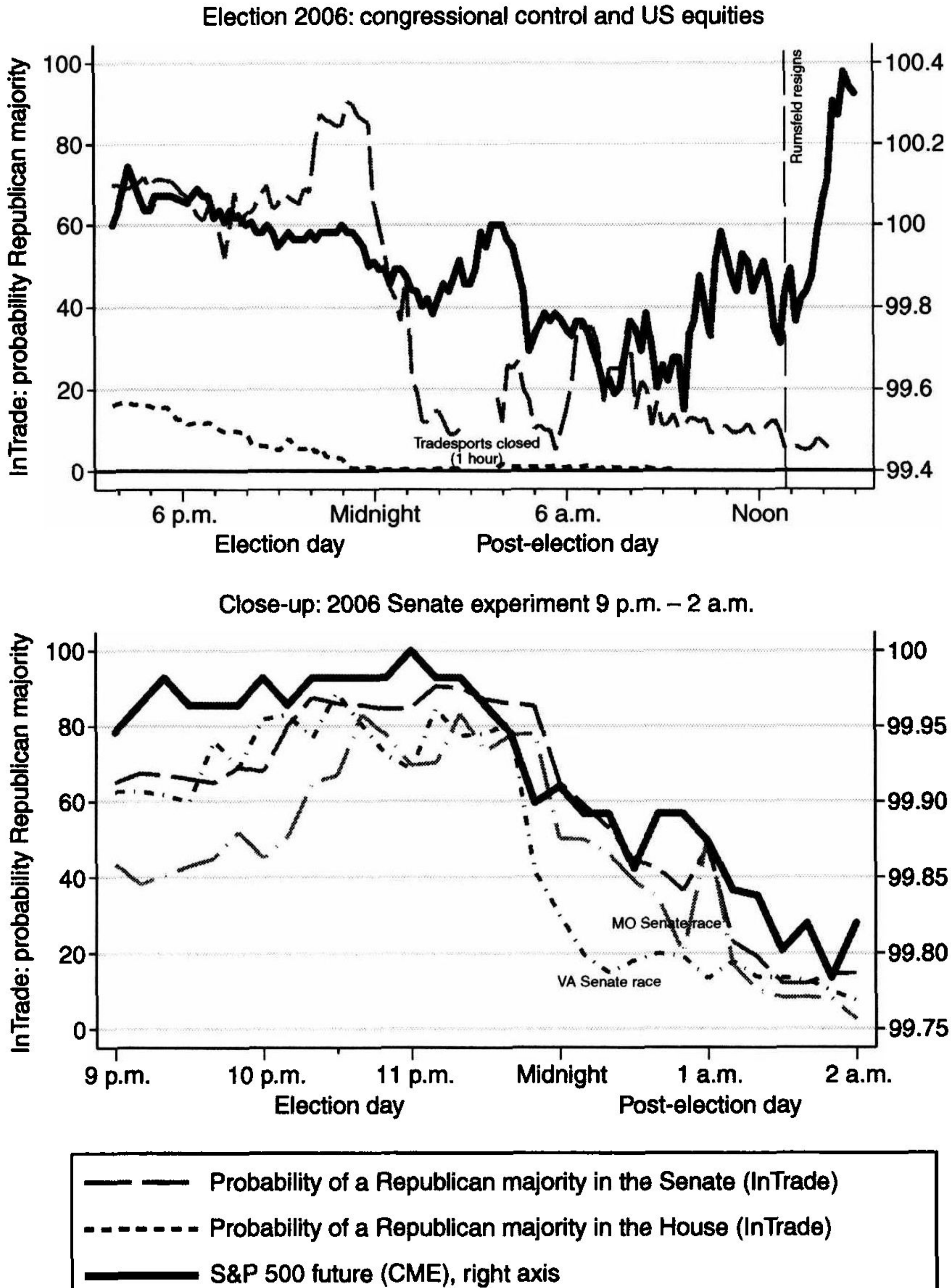


Figure 3.3 Prediction markets separate equity movements from irrelevant movements.

power of prediction markets to parcel out the market response to election news from changes in the index due to other factors.[12] Moreover, this parceling reduces noise, allowing more precise estimates. To see this, compare the standard errors in Table 3.1 to the standard error of 0.5 percent of the prediction market event study over the same period.

5 Misuse of prediction markets

Each prediction market contact is a new variable. A new variable will have correlations with other variables, and thus chances to mistake correlation for causation. To date, the most common instance of such confusion involving prediction markets comes from studies that regress stocks and stock market indices on prediction markets in the weeks and months before an election.

Figure 3.4 presents a stylized, but relatively complete, representation of the system relating economic and political events, voters, prediction markets and financial markets. Researchers have generally been interested (as we are in this chapter) in discovering the effect of the electoral fortunes of various candidates on financial markets. This effect is represented by the large arrow in the figure.

Estimating the effect of politicians on the market is complicated by the fact that the health of the economy in general, and financial markets in particular, have a profound impact on the (re-)election chances of politicians.[13] A simple regression of the stock market on prediction market prices will be biased by reverse causality. For example, Herron (2000), which studies the relationship between Labour's fortunes in the 1992 British parliamentary elections and the FTSE 100, the standard British stock index, speaks about this issue, but is forced to assume that reverse causality is not an issue due to data limitations. Reverse causality is likely here, so the results of this study are likely biased.[14]

A straightforward illustration of this bias is found in Figure 3.5, which shows the S&P 500 and Bush's probability of re-election over the course of the 2004 election. Note that the variables are initially uncorrelated, and then become correlated as the election nears its conclusion. This mirrors the finding in Fair (1996), that economic performance in the two quarters preceding a presidential election are important predictors of presidential vote shares.

Regressing the S&P 500 on the probability of Bush's re-election over the entire time frame of Figure 3.5 yields estimates of Bush's impact on the S&P

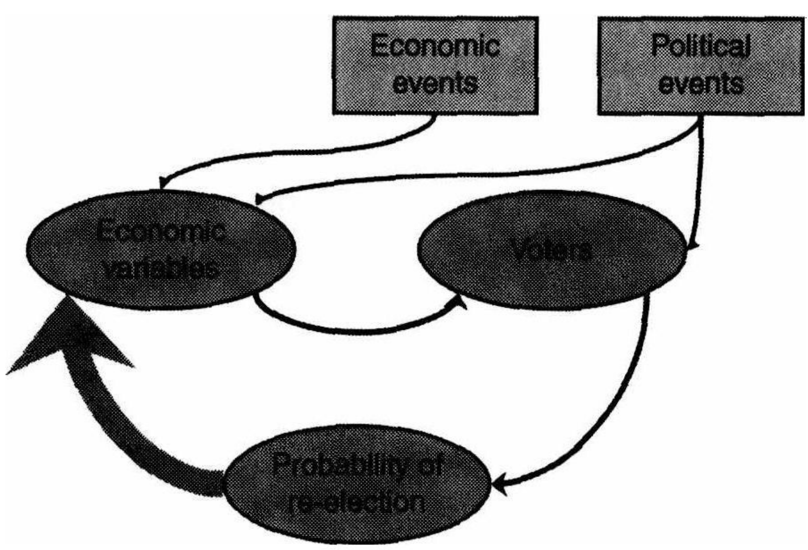


Figure 3.4 The interactions between information, voters, prediction markets and financial markets are complicated.

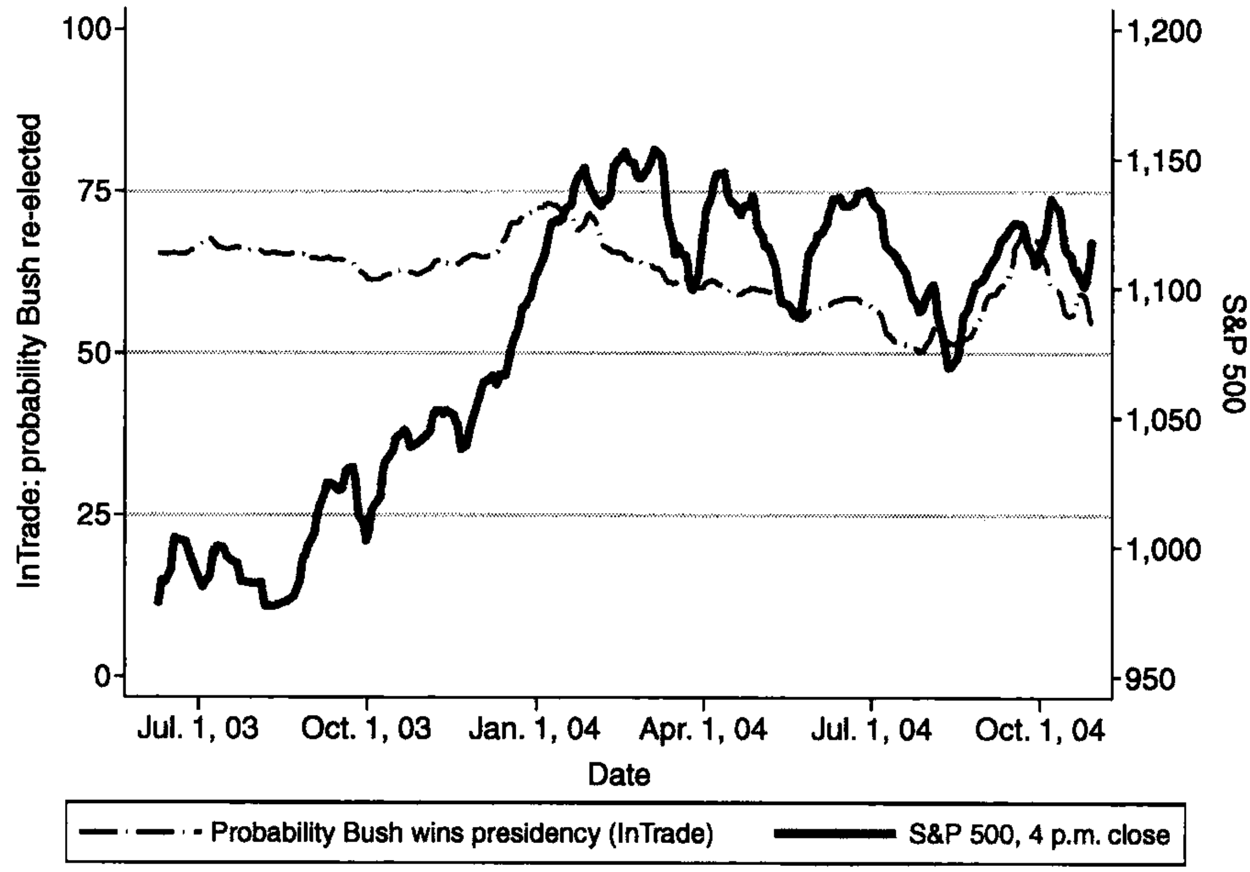


Figure 3.5 Bush's chances of re-election and the S&P 500 become increasingly correlated over time.

Note
Graph displays a five-day moving average of the closing prices at 4 p.m. EST for both the S&P 500 and the InTrade prediction market contract on Bush's re-election.

500 (vs. Kerry) that are as high as 24 percent. This stands in stark contrast to the actual effect of Bush on the S&P 500 of only 2 percent.

Recognizing the possibility for such bias studies such as Herron *et al.* (1999), Knight (2006) and Mattozzi (2008) instead regress industry stock portfolios on prediction market prices. It is unlikely that the economic performance of a single industry would have a significant impact on the probability of Bush's or Kerry's election, side-stepping concerns about reverse causality. However, such studies may present other biases, also illustrated in Figure 3.4.

Specifically, Figure 3.4 shows that other factors may affect both an industry's performance and a candidate's fortunes. This is best illustrated with an example taken from Snowberg *et al.* (2007a). Suppose that an election features a pro- and anti-war candidate, and the pro-war candidate is a more capable war president. If shares in defense contractors increase in value when the pro-war candidate's electoral prospects improve, one might be tempted to conclude that the defense contractors stocks are worth more because there is a higher chance that the pro-war candidate will be elected. However, a third factor – such as threatening actions from another nation – may have led both numbers to appreciate: the defense contractor's from their increased sales in an increasingly likely war, and the pro-war candidate's from his country's increased need of his leadership in wartime.

Such pre-election analyses are popular with political commentators and the press. Indeed, in July 2008, a month after Obama had clinched the Democrat nomination (and four months before the general election) commentators noted the correlation between Obama's increasing chance of winning the presidential election and the declining stock market illustrated in Figure 3.6. The correlation between the data series in Figure 3.6 is slightly more than –0.9. Some commentators even went so far as to describe this pattern of equity prices as the "Obama Slump."

Given the evidence in this chapter and our previous research, it is far more likely that a declining economy was hurting both the stock market and McCain's chance of election. This did not prevent outlets such as the *Wall Street Journal* from publishing such speculation on their website.[15]

As with all endogeneity concerns, the remedy is some form of instrumental variable or a natural experiment. While it is theoretically possible to construct a prediction market that would be correlated with one variable and uncorrelated with another in the analysis, and thus be able to serve as an instrumental variable (Wolfers and Zitzewitz, 2005), to date we are unaware of a successful application of this procedure. Instead, we focus on *instrumental events* such as debates or gaffes (like the infamous "Dean Scream") that affect the political fortunes of the candidates without directly affecting the economy. By focusing on election night in this chapter, we examine a period when votes have already been cast,

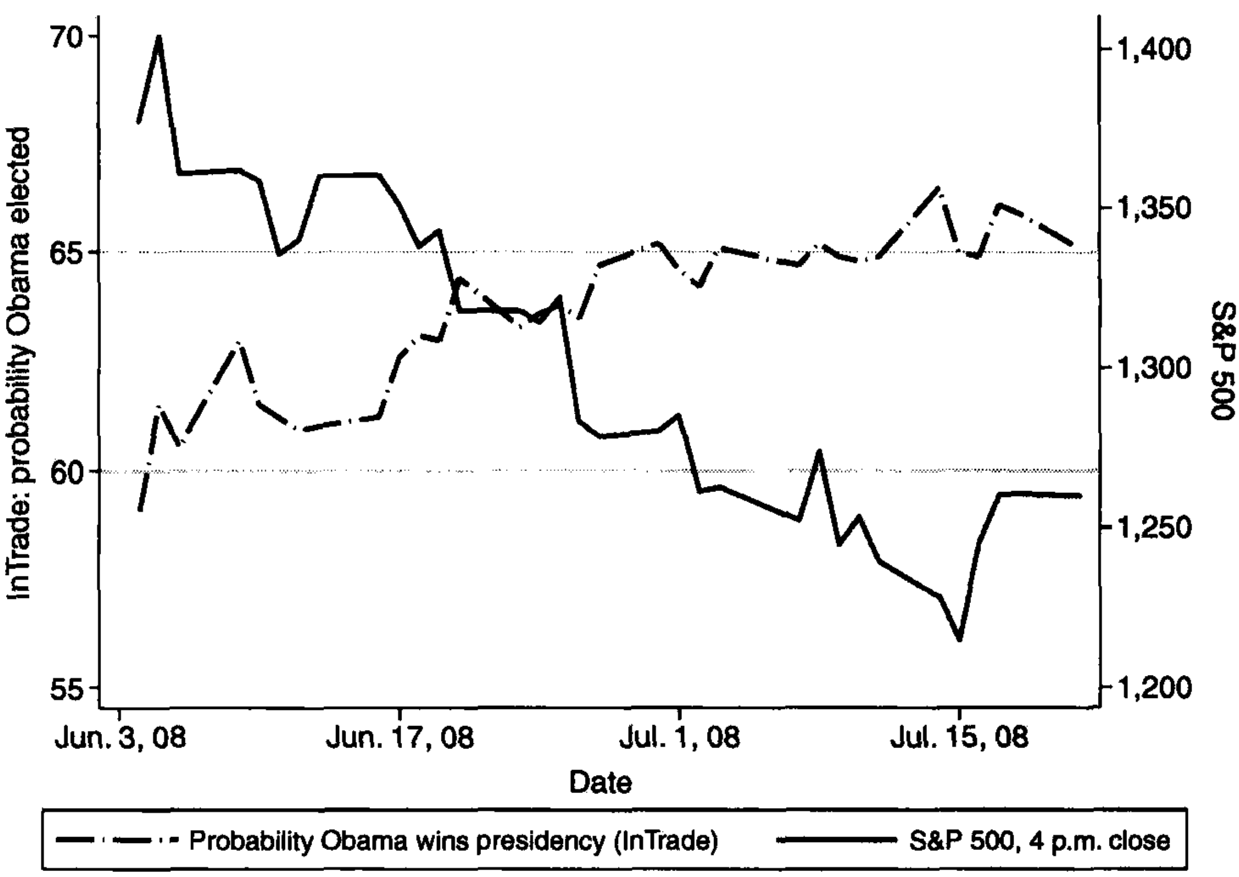


Figure 3.6 The high negative correlation between equities and Obama's chance of winning are likely due to poor economic performance hurting McCain.

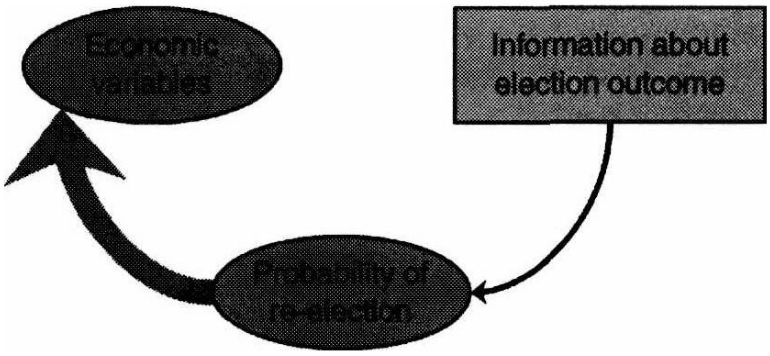


Figure 3.7 On election night, information can only affect economic variables through the expectations about each candidate's chances of election.

and thus the information that is revealed as votes are counted can only affect the economy through changes in expectations, not through changes in vote choice as illustrated in Figure 3.7.

6 Conclusion

While we have documented problems that plague many event studies, we do not mean to imply that all event studies without prediction markets will be flawed. For example, reverse causality may not be a problem when the economic object of study is unimportant for electoral outcomes, such as in Slemrod and Greimel (1999), which studies municipal bond markets. To take another recent example, Jayachandran (2006) studies the market response to different majority parties in the Senate by using an instrumental event, the switch of Senator Jeffords from the Republican party to an independent, which handed control of the Senate to the Democrats. To conduct her study, she created portfolios of stocks associated with Democrats and Republicans, and then observed changes in market value as news about Jeffords' switch leaked out over the span of a week. Her study shows clearly that companies associated with Democrats increased in value relative to those associated with Republicans. What she could not study was what happened to the market as a whole, as many other events would have increased or decreased the value of a broad market index like the S&P 500. So while Democrat stocks increased slightly in value relative to Republican stocks, her study leaves open the possibility that all stocks rose (or fell) substantially due to Jeffords' switch.

Prediction markets erase this constraint. As we have shown in this chapter, broad market indices move very little in response to a switch in the majority party of the Senate. By augmenting event studies with prediction markets, other scholars will no doubt come up with creative ways to address many other unanswered questions.

Notes

1 Event studies and the analysis of prediction markets rely on the efficient-markets hypothesis, namely that all available information is quickly reflected in the price of financial securities. For some problems with this hypothesis, see Malkiel (2003). Regardless of these potential problems, empirically speaking, prediction markets have proven to be extremely accurate (Wolfers and Zitzewitz, 2004; Berg *et al.*, 2008).

2 This is a simplified version of the full statistical methodology. For a complete treatment, see Schwert (1981).

3 For a summary of different types of prediction markets, see Wolfers and Zitzewitz (2004).

4 Wolfers and Zitzewitz (2006) show that under reasonable assumptions of trader risk-attitudes, prediction market prices accurately reflect underlying probabilities.

5 This research was not the first to use prediction markets to examine the impacts of politics on economic variables. Earlier examples include Slemrod and Greimel (1999), Herron (2000), Leigh *et al.* (2003) and Wolfers and Zitzewitz (2009).

6 At the time of the 2004 election, InTrade was known as TradeSports. We use the company's current name to lessen confusion. As US stock exchanges are closed during the evening and night, we use an S&P future with a near-month delivery date rather than the actual S&P 500.

7 More technically, we estimated a first differences model of the form:

$$\Delta(\text{S\&P 500})_t = \alpha_t + \beta_t \Delta\,(\text{Prob: Bush re-election})_t + \varepsilon_t$$

where differences were taken over 30-minute periods. In order to correct for heteroskedacity introduced by the occasional missing observation, we weighted by one over the square root of the number of periods since the last observation. We also use White (1980) heteroskedactic consistent standard errors. The estimates produced via this method for the S&P 500 are roughly consistent with other broad market indices, see Snowberg *et al.*, 2007a) for details. Readers may be concerned that these results may mask large sector increases or decreases in stock market value. Dooley (2008) shows this was not the case, at least in 2004.

8 Note that even after the networks had called Ohio for Bush, the prediction market hovered between a 90 percent and 95 percent chance of a Bush victory, suggesting that the markets might have learned something about the uncertainty of election results from the 2000 recount. A traditional event study could not take account of this nuance. Note also that the last two possible closes of the event window are especially focal if researchers are using only actual market prices rather than the prices from futures markets.

9 Until 1984 the market was not open on election day.

10 InTrade ran a contract on the probability of Rumsfeld's resignation. Although thinly traded, it suggests that his resignation led to a 1.5 percent increase in the S&P 500. It should be noted that this rise may not have been caused entirely by Rumsfeld's resignation, but may be a response to some sign of a change in the Bush administration's policy in Iraq and the Middle East more broadly.

11 A similar study in 2002 showed that the S&P 500 increased by 0.6 percent in response to a shift from a Democrat to a Republican Senate. See Snowberg *et al.* (2007b).

12 Moreover, the results do not change if the contracts for the House and Senate are entered together or separately, indicating that this is not the result of a joint effect of the loss of both Houses of Congress.

13 The academic literature on economic voting (Kramer, 1971; Fiorina, 1981; Lewis-Beck, 1988) is truly staggering. See Lewis-Beck and Paldam (2000) for a review. Particularly relevant here is the theory of sociotropic (Kinder and Kiewiet, 1979, 1981) voting, which is based on the observation that the perception of general

economic variables such as the stock market is strongly related to vote choice, whereas personal economic experience is less strongly correlated.

14 Note that classical event studies will not always be biased by reverse causality. For example, in Slemrod and Greimel (1999) it is unlikely that the municipal bond market had any effect on the outcome of the Republican presidential primary, and in Wolfers and Zitzewitz (2009) it is unlikely that the price of oil changed the probability that the United States invaded Iraq.

15 See http://online.wsj.com/article/SB121592969771748931.html and midasoracle.org/2008/07/15/intrade/dow-jones/.

References

Berg, Joyce, Robert Forsythe, Forrest Nelson and Thomas Rietz. 2008. "Results from a Dozen Years of Election Futures Markets Research." In *The Handbook of Experimental Economics Results*, ed. Charles R. Plott and Vernon L. Smith. Elsevier: Handbooks in Economics series.

Den Hartog, Chris and Nathan W. Monroe. 2008. "The Value of Majority Status: The Effect of Jeffords' Switch on Asset Prices of Republican and Democratic Firms." *Legislative Studies Quarterly* 33(1):63–84.

Dooley, Patrick. 2008. "Elections and Market Capitalization Effects: Evidence from the 2004 US Presidential Election." Dartmouth College, mimeo.

Downs, Anthony. 1957. *An Economic Theory of Democracy*. New York: Harper Collins.

Fair, Ray C. 1978. "The Effect of Economic Events on Votes for President." *The Review of Economics and Statistics* 60(2):159–173.

Fair, Ray C. 1996. "Econometrics and Presidential Elections." *The Journal of Economic Perspectives* 10(3):89–102.

Fair, Ray C. 2004. "Presidential Vote Equation – October 29, 2004." Yale University, mimeo, http://fairmodel.econ.yale.edu/vote2004/vot1004.htm.

Fiorina, Morris P. 1981. *Retrospective Voting in American National Elections*. New Haven, CT: Yale University Press.

Fisman, Ray. 2001. "Estimating the Value of Political Connections." *The American Economic Review* 91(4):1095–1102.

Gilligan, Thomas W. and Keith Krehbiel. 1988. "Complex Rules and Congressional Outcomes: An Event Study of Energy Tax Legislation." *Journal of Politics* 50(3):625–654.

Herron, Michael C. 2000. "Estimating the Economic Impact of Political Party Competition in the 1992 British Election." *American Journal of Political Science* 44(2):326–337.

Herron, Michael C., James Lavin, Donald Cram and Jay Silver. 1999. "Measurement of Political Effects in the United States Economy: A Study of the 1992 Presidential Election." *Economics and Politics* 11(1):51–81.

Jayachandran, Seema. 2006. "The Jeffords Effect." *The Journal of Law and Economics* 49(2):397–425.

Kinder, Donald R. and D. Roderick Kiewiet. 1979. "Economic Discontent and Political Behavior: The Role of Personal Grievances and Collective Economic Judgments in Congressional Voting." *American Journal of Political Science* 23(3):495–527.

Kinder, Donald R. and D. Roderick Kiewiet. 1981. "Sociotropic Politics: The American Case." *British Journal of Political Science* 11(2):129–161.

Knight, Brian. 2006. "Are Policy Platforms Capitalized into Equity Prices? Evidence from the Bush/Gore 2000 Presidential Election." *Journal of Public Economics* 90(4–5):751–773.

Kramer, Gerald H. 1971. "Short-Term Fluctuations in US Voting Behavior, 1896–1964." *The American Political Science Review* 65(1):131–143.

Leigh, Andrew, Justin Wolfers and Eric Zitzewitz. 2003. "What do Financial Markets Think of the War with Iraq?" *NBER Working Paper #9587.*

Lewis-Beck, Michael S. 1988. *Economics and Elections: The Major Western Democracies.* Ann Arbor: University of Michigan Press.

Lewis-Beck, Michael S. and Martin Paldam. 2000. "Economic Voting: An Introduction." *Electoral Studies* 19(2–3):113–121.

Malkiel, Burton G. 2003. "The Efficient Market Hypothesis and Its Critics." *Journal of Economic Perspectives* 17(1):59–82.

Mattozzi, Andrea. 2008. "Can We Insure Against Political Uncertainty? Evidence from the U.S. Stock Market." *Public Choice* 137(1–2):43–55.

Monroe, Nathan W. 2008. "Majority Party Dividends: The Effect of the 1994 Election on Asset Prices of Republican Constituent Firms." Michigan State University, mimeo.

Rhode, Paul W. and Koleman S. Strumpf. 2004. "Historical Presidential Betting Markets." *Journal of Economic Perspectives* 18(2):127–142.

Rhode, Paul W. and Koleman S. Strumpf. 2006. "Manipulating Political Stock Markets: A Field Experiment and a Century of Observational Data." University of Arizona, mimeo.

Rhode, Paul W. and Koleman S. Strumpf. 2008. "Historical Election Betting Markets: An International Perspective." University of Arizona, mimeo.

Roberts, Brian E. 1990a. "A Dead Senator Tells No Lies: Seniority and the Distribution of Federal Benefits." *American Journal of Political Science* 34(1):31–58.

Roberts, Brian E. 1990b. "Political Institutions, Policy Expectations, and the 1980 Election: A Financial Market Perspective." *American Journal of Political Science* 34(2):289–310.

Rogerson, William P. 1989. "Profit Regulation of Defense Contractors and Prizes for Innovation." *The Journal of Political Economy* 97(6):1284–1305.

Santa-Clara, Pedro and Rossen Valkanov. 2003. "The Presidential Puzzle: Political Cycles and the Stock Market." *The Journal of Finance* 58(5):1841–1872.

Schnietz, Karen E. 2000. "The Institutional Foundation of US Trade Policy: Revisiting Explanations for the 1934 Reciprocal Trade Agreements Act." *Journal of Policy History* 12(4):417–444.

Schwert, G. William. 1981. "Using Financial Data to Measure Effects of Regulation." *Journal of Law and Economics* 24(1):121–158.

Slemrod, Joel and Timothy Greimel. 1999. "Did Steve Forbes scare the US Municipal Bond Market?" *Journal of Public Economics* 74(1):81–96.

Snowberg, Erik, Justin Wolfers and Eric Zitzewitz. 2007a. "Partisan Impacts on the Economy: Evidence from Prediction Markets and Close Elections." *The Quarterly Journal of Economics* 122(2):807–829.

Snowberg, Erik, Justin Wolfers and Eric Zitzewitz. 2007b. "Party Influence in Congress and the Economy." *The Quarterly Journal of Political Science* 2(3):277–286.

White, Halbert. 1980. "A Heteroskedasticity-Consistent Covariance Matrix Estimator and a Direct Test for Heteroskedasticity." *Econometrica* 48(4):817–838.

Wolfers, Justin and Eric Zitzewitz. 2004. "Prediction Markets." *Journal of Economic Perspectives* 18(2):107–126.

Wolfers, Justin and Eric Zitzewitz. 2005. "Five Open Questions About Prediction Markets." In *Information Markets: A New Way of Making Decisions in the Public and Private Sectors*, ed. Robert Hahn and Paul Tetlock. Washington, DC: AEI-Brookings Press.

Wolfers, Justin and Eric Zitzewitz. 2006. "Interpreting Prediction Market Prices as Probabilities." *NBER Working Paper #12200.*

Wolfers, Justin and Eric Zitzewitz. 2009. "Using Markets to Inform Policy: The Case of the Iraq War." *Economica* 76(302):225–250.

4 Mechanisms for prediction markets

Yiling Chen

A prediction market offers contingent contracts whose future payoff is tied with the outcome of some uncertain event of interest and attracts traders to wager on the outcome. For example, to predict how likely it is that there will be a bird flu outbreak in the United States by 2012, a prediction market can offer a contract that pays \$1 if a bird flu case is confirmed in United States by the end of 2012, and \$0 otherwise. A risk-neutral agent who believes that a bird flu outbreak will happen with probability α has the incentive to trade the contract and drive the price of the contract to $\$\alpha$. If every participant trades the contract based on his or her private information, at an equilibrium the market price of the contract can represent the consensus belief on the likelihood of a bird flu outbreak based on the pooled information. In theory, the event of interest can be thought of as a random variable X, and the payoff of the contract depends on the realized value of X. X is often a discrete or discretized random variable that has n mutually exclusive and exhaustive outcomes.

Different market mechanisms can be used to facilitate trading in prediction markets. As the primary function of prediction markets is information aggregation, the following properties, among others, are desirable for market mechanisms to better achieve the information aggregation goal:

- *Liquidity.* Liquidity requires that market participants can find their counterparties to trade whenever they want. Participants reveal their information by trading contracts. If they cannot trade, even if they have relevant information, they cannot reveal it in the market.
- *Expressiveness.* Expressiveness refers to giving market participants more freedom to express their information. Typically, this is achieved by defining more expressive betting languages that participants can use.
- *Bounded budget.* Bounded budget restricts the potential loss of the market institution.

In this chapter, we introduce market mechanisms that can be used to operate prediction markets, and discuss them in the context of these properties. For all market mechanisms, we assume that the market institution does not take any transaction fee or commission. This is certainly not true in reality, but it allows us to better explain the mechanisms. In addition, because prediction markets aim to elicit and aggregate information, under many situations it is reasonable for the market institution to not make a profit or even subsidize the market, because it can benefit from the resultant information.

1 Auctioneer mechanisms

Auctioneer mechanisms refer to those markets where the market institution risklessly matches buy and sell orders of contingent contracts. Because the market institution plays a role of order-matching and does not incur any loss no matter what the realized values of the contingent contracts are, they are often called an auctioneer. Clearly, auctioneer mechanisms have bounded budgets.

1.1 Call markets

Call markets[1] are the mechanism used by stock exchanges in their early years. In a call market, individual contracts are traded at pre-specified times. Participants submit limit orders. A limit order i can be represented by a triple (φ_i, q_i, b_i), where φ_i is the contract to trade, q_i is the number of shares to buy, with a negative value denoting a sell order, and b_i is the limit price, which is the maximum price per share the participant is willing to pay for a buy order and the minimum price per share the participant is willing to accept for a sell order. b_i is often called the *bid price* if i is a buy order and the *ask price* if i is a sell order. All orders for the same contract are assembled at a pre-specified time to determine a market clearing price at which the demand to buy the contract equals the supply to sell the contract. Buy orders whose bid prices are higher than or equal to the clearing price and sell orders whose ask prices are lower than the clearing price are accepted. All transactions happen at the market clearing price.

We use an example to illustrate how a call market works. Suppose we have a single contract that pays $1 if and only if event A happens. We denote it contract A.

Example 1: The call market receives the following nine orders in sequence before the pre-specified clearing time:

1		(A, 1, $0.28)
2	(A, –1, $0.30)	
3		(A, –1, $0.24)
4	(A, 1, $0.10)	
5		(A, –1, $0.08)
6	(A, 1, $0.20)	
7		(A, –1, $0.16)
8	(A, 1, $0.06)	
9		(A, –1, $0.32)

Table 4.1 shows the demand and supply of contract A at different price ranges.

Table 4.1 Demand and supply of contract A at different price ranges

Price	(−∞, 0.06]	(0.06, 0.08)	[0.08, 0.10]	(0.10, 0.16)	**[0.16, 0.20]**	(0.20, 0.24)	[0.24, 0.28]	(0.28, 0.30)	[0.30, 0.32)	[0.32, ∞)
Demand	4	3	3	2	**2**	1	1	0	0	0
Supply	0	0	1	1	**2**	2	3	3	4	5

Demand equals supply at any price within the closed interval [0.16, 0.20]. Different rules can be used to decide the exact clearing price. For this example, we use the mid-point of the interval, \$0.18, as the clearing price. Orders 1, 5, 6, 7 are accepted and traded at the price of \$0.18. The other orders are rejected or left in the pool for the next matching.

In essence, a call market is a *k-double auction*.[2] Suppose the call market receives N buy orders and M sell orders.[3] One way to determine the price is to use an *M-th price auction*. A M-th price auction ranks all $N+M$ orders in decreasing order of limit prices. The market price is set to be the M-th highest limit price, p^M. If there are x buy orders in the top M orders, there must be x sell orders in the bottom N orders. The x buy orders have limit prices higher than or equal to p^M, and the x sell orders have limit prices lower than p^M. At p^M, both demand and supply equal x. The market clears at p^M. In Example 1, M equals 5 and p^M equals \$0.20. Another way to determine the price is to use an *$(M+1)$-th price auction*. In such an auction, the market price is set to be the $(M+1)$-th highest limit price, p^{M+1}. This price also guarantees market clearance. In Example 1, p^{M+1} is \$0.16. The k-double auction sets the market clearing price to be $p^{M+1}+k(p^M-p^{M+1})$, where $k\in[0, 1]$. The 0-double auction is the $(M+1)$-th price auction, while the 1-double auction is the M-th price auction. A 0.5-double auction will set the market clearing price in Example 1 to \$0.18.

1.2 Continuous double auctions

Continuous double auctions (CDA)[4] are the mechanism currently used by stock exchanges. CDA is a k-double auction repeated whenever an order comes in. Unlike call markets, where orders are processed in batches, CDA matches orders in real-time – as soon as the bid price of a buy order is higher than the ask price of a sell order, a transaction happens. The auctioneer keeps an order book. At any time, unmatched buy orders are listed in decreasing order of their bid prices and unmatched sell orders are listed in increasing order of their ask prices. The highest bid is lower than the lowest ask in the order book. This difference is called the bid–ask spread. When a new order arrives, a transaction happens immediately if the new order or part of it can be matched with orders in the order book. The unmatched part is left in the order book for future matching.

Consider the nine orders in Example 1. Suppose a CDA is used, and the price is set by a 0.5-double auction whenever there is a match. In other words, whenever the bid price of a buy order is higher than the ask price of a sell order, a transaction happens at the mid-point of the bid and ask prices. Orders 1 and 2 sit in the order book when they arrive to the market because they can not be matched with each other. The bid–ask spread after receiving order 2 is \$0.02. When order 3 arrives, it is immediately matched with order 1 and the transaction price is \$0.26. Similarly, order 5 is matched with order 4 at a price of \$0.09, and order 7 is matched with order 6 at a price of \$0.18.

CDA is used by most prediction markets, including Iowa Electronic Markets (IEM),[5] Intrade,[6] Newsfutures[7] and the internal prediction markets of HP[8] and Google.[9]

Liquidity can be a concern for CDA at times. When there are not enough traders, CDA is subject to the *thin market* problem. A participant who wants to trade may not be able to find a counter-party to trade with. The market prices are set by relatively few transactions in a thin market and cannot represent the true market conditions. Hence, information aggregation may be put into question. This is especially a problem for many prediction markets because they have far less traders than most financial markets.

Because call markets process orders in batches, they can offer better liquidity than CDA. The market price in a call market is more stable than that in a CDA for a thinner market. However, the increased liquidity in a call market is at the expense of delayed transactions. New information cannot be reflected in the market price in a call market immediately. Many stock exchanges have adopted call markets for daily opening and closing of the stock markets to obtain more stable and informative opening and closing prices, and use CDA the rest of the time.

1.3 Generalized call mechanisms and combinatorial prediction markets

Both call markets and CDA are bilateral – matching buy and sell orders of the same contract. However, call markets can be generalized to perform multilateral matching. *Combined value trading*[10] and *pari-mutuel call markets*[11] are multilateral call markets designed to reduce execution risk[12] and increase liquidity for financial markets. The ability to perform multilateral order matching makes generalized call mechanisms a suitable option for operating combinatorial prediction markets.[13]

Consider a US presidential election. If we are interested in predicting which of the two political parties, the Democratic Party or the Republican Party, will win the election in which states, we can operate 50 independent markets, one for each state. Each market offers two contracts, one for each political party; paying off $1 if the corresponding political party wins the election in the corresponding state. However, if a trader has information on the likelihood that the Democratic Party will win both Florida and Ohio but not New York, such information cannot be fully expressed in these independent markets. To increase expressiveness, a combinatorial prediction market considers all elections together. We have 50 binary base events, i.e., elections in 50 states. The outcome space thus consists of all possible complete specifications of the 50 election results. The size of the outcome space is 2^{50}. It is typically not interesting to offer one contract for each outcome. Instead, a combinatorial prediction market defines some expressive betting language and allows participants to bet on combinations of outcomes according to the betting language. For example, a two-clause Boolean betting language allows participants to specify the conjunction or disjunction of outcomes of two base events and create a contract that pays off $1 if the Boolean formula turns out to be true. Because participants can submit limit orders for different contracts in a combinatorial prediction market, bilateral matching cannot provide enough liquidity. The auctioneer sometimes can safely perform multilateral matching without incurring any risk. For instance, the three orders (Democratic Party wins Florida and Ohio, 1, $0.4), (Democratic Party loses Ohio and wins New York, 1, $0.5), and (Democratic Party loses Florida and New York, 1, $0.3) can be accepted simultaneously at their limit price, because the auctioneer can collect $1.2 now but needs to pay out $1 for, at most, one contract in the future.

The order matching problem of a combinatorial prediction market can be modeled as an optimization problem. If every combinatorial contract pays off $1 if and only if the specified event happens, an auctioneer hoping to maximize its worst-case profit can determine what orders to accept by solving the following linear integer program,

$$\max_{x_i c} \quad c \tag{4.1}$$

$$\text{s.t. } \Sigma_i (b_i - I_i(w))q_i x_i \geq c_i \;\; \forall w \in \Omega$$
$$x_i \in \{0,1\}, \qquad \forall i \in O$$

Ω is the outcome space. O is the set of received orders. For each order i, q_i is the quantity to buy (negative denotes sell); b_i is the limit price; $I_i(\mathrm{w})$ is an indicator variable that equals 1 when the event specified in order i is true and 0 otherwise. x_i is the binary decision variable that equals 1 when the auctioneer accepts order i and 0 otherwise. c is the worst-case profit of the auctioneer, which is at least 0 because the auctioneer can simply reject all orders. Accepted orders are traded at their limit price. The auctioneer can have other objectives such as maximizing the total trades. If partial orders are allowed, x_i can take a real value in [0, 1], representing the fraction of order i that the auctioneer accepts. An optimal solution to the dual problem, which in general is not unique, gives the market clearing prices for w. The Convex Pari-mutuel Call Auction Mechanism (CPCAM)[14] introduces a regularization term to the objective function to ensure the uniqueness of the market clearing prices.

Attentive readers might have noticed that the outcome space Ω for the presidential election example is exponentially large. Thus, the optimization problem (4.1) has a large number of constraints and may not be solved efficiently in the worst case. Researchers have been studying the computational complexity of operating combinatorial call market mechanisms for Boolean and permutation combinatorics. For Boolean combinatorics as in the presidential election example, if participants can bet on any Boolean formula of the base events, the auctioneer's order matching problem is computationally hard.[15] For permutation combinatorics, the outcome space consists of all possible rank orders of competing candidates, a combinatorial prediction market allows participants to bet on properties of the final ordering. A call mechanism can efficiently handle *subset betting* where a participant can bet on a particular candidate taking one of a subset of positions (e.g., horse A finishes at the first or third position) or one of a subset of candidates taking one position (e.g., horses A or B will finish at the first position) if the auctioneer can accept partial orders.[16] When partial orders are not acceptable, the auctioneer's order matching problem can be solved efficiently for subset betting only when the betting language is restricted to one candidate and one position.[17] Agrawal *et al.*[18] studied permutation betting in CPCAM. They allowed participants to bet on linear combinations of candidate–position specifications and showed that the auctioneer's order matching problem can be solved efficiently by a convex optimization problem. Betting on the relative positions of two candidates (e.g. horse A beats horse B) is computationally hard to support in a call mechanism.[19]

2 Pari-mutuel markets

Pari-mutuel markets are the mechanism often used for horse races.[20] In a pari-mutuel market, participants place wagers on one of two or more mutually exclusive and exhaustive outcomes of some event. After the true outcome becomes known, all money that is bet on incorrect outcomes is redistributed to those who bet on the correct outcomes, in proportion to the amount they wagered. Let W_i represent the total amount of money wagered on outcome i. $W=\Sigma_i W_i$ is the total money wagered in the market. If outcome j happens, a participant who bets on the outcome 5 gets $\frac{W}{W_j}$ dollars for every dollar he wagers. A participant can always bet in a pari-mutuel market as long as he believes it to be profitable. Hence, pari-mutuel markets have the advantage of infinite liquidity.

Unlike auctioneer mechanisms, which offer contracts whose future payoff under every outcome is fixed, pari-mutuel markets do not have an explicit notion of contracts. Implicitly, participants pay \$1 to get one share of a contract that entitles its owner to an equal share of the total money wagered if the corresponding outcome happens. The payoff of the contract is not fixed at the time of betting – it is not determined until the pool is closed. Participants in pari-mutuel markets may have incentives to place their bet at the last minute, because the price of the contract remains the same, but placing a bet early reveals a participant's information to others.

3 Automated market maker mechanisms

In auctioneer mechanisms and pari-mutuel markets, market participants play a zero-sum game – the net profit of all participants is always zero. These mechanisms face the theoretical challenge of *no-trade theorems*,[21] which imply that rational risk-neutral agents should not trade in zero-sum markets. The intuition of no-trade theorems is that if market participants are rational, have common prior and rationality is common knowledge, a participant's intention to trade indicates that he can benefit from the transaction. His trading partner should then reason that he must know something that she doesn't know, and should revise her belief and reject the trade.

In an automated market maker mechanism, the market maker, which is the market institution, sets prices for buy and sell orders. All transactions are with the market maker. The market maker bears risk and hence may incur loss. This turns the market into a positive-sum game for market participants. Even rational participants have incentives to trade. Moreover, because a participant can always trade with the market maker whenever he wants, automated market maker mechanisms overcome the thin market problem that the auctioneer mechanisms suffer from.

The bounded budget is an important property for automated market makers. If the loss of a market maker is unbounded, the market mechanism can hardly be practical. In this section, we discuss several automated market maker mechanisms that have bounded budgets.

3.1 Market scoring rules and cost function-based market makers

Market scoring rules have become the de facto market maker mechanism for prediction markets. They are based on proper scoring rules. We hence start with introducing proper scoring rules.

Proper scoring rules

Without markets, a natural way to get information about an event is to directly ask experts for their probability assessments of the event. But if experts get no reward for providing accurate information, they may not be truthful. *Proper scoring rules* are rewarding functions that are used to incentivize experts to report their probability assessments truthfully. Formally, let v represent a discrete random variable to be predicted, with n mutually exclusive and exhaustive outcomes. Let $\mathbf{r}=\langle r_1, r_2,\ldots, r_n\rangle$ be a reported probability estimate for the random variable v. A *scoring rule* is a sequence of scoring functions, $S=\{s_1(\mathbf{r}), s_2(\mathbf{r}),\ldots, s_n(\mathbf{r})\}$, such that a score $s_i(\mathbf{r})$ is assigned to $\mathbf{r}$ if outcome i of the random variable v is realized. A scoring rule is *(strictly) proper* if a risk-neutral expert (strictly) maximizes his expected score by reporting truthfully. In other words, proper scoring rules are *incentive compatible* for risk-neutral experts when eliciting probability assessments. A vast class of scoring rules is proper. Given any bounded (strictly) convex function $G(\mathbf{r})$,

$$s_i(\mathbf{r}) = G(\mathbf{r}) - G'(\mathbf{r})\cdot\mathbf{r} + G_i'(\mathbf{r}), \tag{4.2}$$

where $G'(\mathbf{r})=\langle G_1'(\mathbf{r}), G_2'(\mathbf{r}),\ldots G_n'(\mathbf{r})\rangle$ is the subgradient of $G(\mathbf{r})$, defines a (strictly) proper scoring rule.[22] Two widely used proper scoring rules are:

$$\text{Logarithmic scoring rule } s_i(\mathbf{r}) = a_i + b\log(r_i) \tag{4.3}$$

$$\text{Quadratic scoring rule } s_i(\mathbf{r}) = a_i + 2br_i - b\sum_{j=1}^{n} r_j^2 , \tag{4.4}$$

where a_i and b are free parameters with $b>0$. Proper scoring rules have been extensively studied over the past five decades,[23] validated in experiments,[24] and applied in a variety of domains.[25] In addition to scoring rules for discrete random variables, there are proper scoring rules for eliciting probability densities of continuous random variables.[26] Proper scoring rules can be adapted to directly elicit some properties of a probability distribution, such as mean and quantiles,[27] without eliciting the whole probability distribution. Shared proper scoring rules that reward each expert according to the difference between his forecast and the average of the others are developed to obtain multiple forecasts.[28] They effectively become wagering mechanisms because experts risk losing money in addition to gaining it.

Market scoring rules

Hanson[29] shows how a proper scoring rule can be converted into a market maker mechanism called a *market scoring rule* (MSR). The market maker uses a proper scoring rule, $S=\{s_1(\mathbf{r}),\ldots, s_n(\mathbf{r})\}$, where $\langle r_1,\ldots, r_n\rangle$ is a reported probability estimate for the random variable v with n exclusive and exhaustive outcomes. The market maker starts the market with some initial probability estimate $\mathbf{r}^0$. Conceptually, every participant in the market may change the current probability estimate to a new estimate of his choice at any time as long as he agrees to pay the scoring rule payment associated with the current probability estimate and receive the scoring rule payment associated with the new estimate. If outcome i is realized, a participant that changes the probability estimate from $\mathbf{r}^{\text{old}}$ to $\mathbf{r}^{\text{new}}$ pays $s_i(\mathbf{r}^{\text{old}})$ and receives $s_i(\mathbf{r}^{\text{new}})$. If a participant only participates in the market once, he effectively faces a proper scoring rule as he has no control over $\mathbf{r}^{\text{old}}$. He is incentivized to report his probability estimate truthfully.

Because participants change the probability estimate in sequence, MSR can be thought of as a sequential shared version of the scoring rule. The market maker pays the last participant and receives payment from the first participant, and hence has bounded loss. The worst-case loss of a market maker is then

$$\max_i \sup_{\mathbf{r}\in\Delta_n}(s_i(\mathbf{r}) - s_i(\mathbf{r}^0)) \tag{4.5}$$

where Δ_n is the probability simplex. For a *logarithmic market scoring rule* (LMSR) with the logarithmic scoring function (4.3), the maximum amount the market maker can lose does not exceed $b\log n$. For a *quadratic market scoring rule* (QMSR) with the quadratic scoring function (4.4), the market maker's worst-case loss is bounded by $\frac{(n-1)b}{n}$.

Cost-function-based market makers

The way we describe MSR above does not make them natural markets. Reporting probabilities could be confusing for market participants, and there isn't any contract to trade in the market. Below, we define another class of market makers, *cost-function based market makers,*[30] which are more natural for implementation purposes. We will show that they are equivalent to MSR.

In a cost function-based market, a market maker offers a contract corresponding to each outcome *i*. The contract associated with outcome *i* pays off $1 if *i* happens, and $0 otherwise. Let q_i be the total quantity of contract *i* held by all traders combined, and let $\mathbf{q}$ be the vector of all quantities held. A cost function-based market maker utilizes a *cost function* $C(\mathbf{q})$ that records the total amount of money that traders have spent as a function of the total number of shares held of each contract. A trader who wants to buy or sell any bundle of contracts such that the total number of outstanding shares changes from $\mathbf{q}_{\text{old}}$ to $\mathbf{q}_{\text{new}}$ must pay $C(\mathbf{q}_{\text{new}})-C(\mathbf{q}_{\text{old}})$ dollars. Negative quantities encode sell orders and negative "payments" encode sale proceeds earned by the trader. The instantaneous price of contract *i* is then $\partial C(\mathbf{q})/\partial q_i$. The price is the cost per share for purchasing an infinitesimal quantity.

The cost function is *valid* if $p_i(\mathbf{q}) \geq 0$ for all *i* and $\mathbf{q}$, and $\Sigma_i p_i(\mathbf{q}) = 1$. The price of a contract is never negative in a valid cost function-based market. If the current price of the contract associated with an outcome *i* were negative, a trader could purchase shares of this contract at a guaranteed profit. If the prices summed to something less than (respectively, greater than) 1, then a trader could purchase (respectively, sell) small equal quantities of each contract for a guaranteed profit. Together, these conditions ensure that there are no arbitrage opportunities within the market and the current prices can always be viewed as a probability distribution over the outcome space. Chen and Vaughan[31] characterize sufficient and necessary conditions for the cost function to be valid:

Theorem 1. *A cost function C is valid if and only if it satisfies the following three properties:*

1 Differentiability: *the partial derivatives* $\partial C(\mathbf{q})/\partial q_i$ exist for all $\mathbf{q} \in \mathbb{R}^N$ *and* $i \in \{1, \ldots n\}$.
2 Increasing monotonicity: *for any* $\mathbf{q}$ *and* $\mathbf{q}'$, *if* $\mathbf{q} \geq \mathbf{q}'$, *then* $C(\mathbf{q}) \geq C(\mathbf{q}')$.
3 Positive translation invariance: *for any* $\mathbf{q}$ *and any constant* k, $C(\mathbf{q}+k\mathbf{1}) = C(\mathbf{q}) + k$.

A cost function-based market maker is equivalent to an MSR market maker if, facing the same market probability estimate, a risk-neutral participant obtains exactly the same profit in both markets under every outcome *i* when truthfully revealing his information. If the current market probability estimate is $\mathbf{p}$ and the participant's probability estimate is $\mathbf{p}'$ the participant will change the market probability estimate to $\mathbf{p}'$ in the MSR and will trade contracts to change the market prices to $\mathbf{p}'$ in the cost function-based market maker. Noting that any convex cost function can be represented as $C(\mathbf{q}) = \sup_{\mathbf{p}\in\Delta_n} (\Sigma_{i=1}^{n} p_i q_i - \alpha(\mathbf{p}))$ for a convex function α, Chen and Vaughan[32] characterize the equivalence between MSR with strictly proper scoring rules and convex cost function-based market makers. Given an MSR with a strictly proper and differentiable scoring rule S, the corresponding convex cost function can be derived by setting

$$\alpha(\mathbf{p}) = \sum_{i=1}^{n} p_i s_i(\mathbf{p}) . \tag{4.6}$$

Given a convex cost-function-based market maker with strictly convex and differentiable $\alpha(\mathbf{p})$, the corresponding MSR uses the following scoring rule:

$$s_i(\mathbf{p}) = \alpha(\mathbf{p}) - \sum_{j=1}^{n} \frac{\partial \alpha(\mathbf{p})}{\partial p_j} p_j + \frac{\partial \alpha(\mathbf{p})}{\partial p_i} , \tag{4.7}$$

It can be easily verified that (4.6) and (4.7) define a one-to-one mapping between MSR with strictly proper and differentiable scoring rules and convex cost-function-based markets with strictly convex and differentiable $\alpha(\mathbf{p})$. The pair of markets are equivalent when prices for all outcomes are positive.

In particular, the cost and price functions of LMSR are:

$$C(\mathbf{q}) = b \log \sum_{j=1}^{n} e^{q_j / b} , \tag{4.8}$$

$$p_i(\mathbf{q}) = \frac{e^{q_i / b}}{\sum_{j=1}^{n} e^{q_j b}} . \tag{4.9}$$

Inking Markets[33] and Microsoft's internal prediction market use LMSR in their cost-function-based format. Figure 4.1 plots the price of contract 1 in a

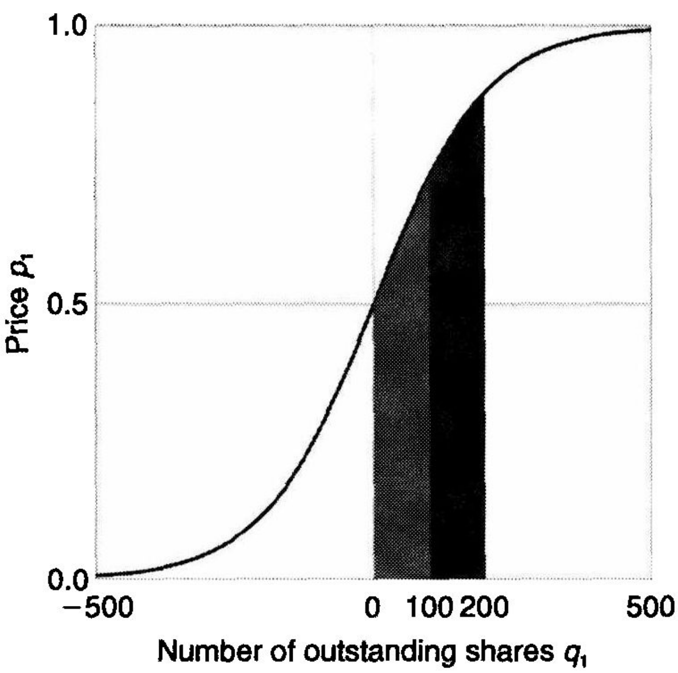


Figure 4.1 Instantaneous price of contract 1 in a two-outcome LMSR. $q_2=0$. $b=100$.

two-outcome LMSR. The curve shows the changes of p_1 when q_1 changes. When the first participant purchases 100 shares of contract 1, the light area captures the amount he needs to pay to the market maker. The second participant's payment for another 100 shares of contract 1 equals the area of the dark region.

Two other families of market maker mechanisms have been proposed based on different rationales. Chen and Pennock[34] introduce *utility-based market makers*. A utility-based market maker has a utility function and always sets the market prices as its risk-neutral probabilities. Agrawal *et al.*[35] propose the *sequential convex pari-mutuel mechanism* (SCPM). It is designed for limit orders. The market maker decides what orders to accept by solving a convex optimization problem. But the payment of the accepted trade is determined by a generalized VCG mechanism. Both the utility-based market makers and SCPM have some equivalence relationships with MSR and cost-function-based market makers.

In MSR and cost-function-based market makers, participants can trade bundles of contracts. Hence, such market makers are natural options for operating combinatorial prediction markets. Researchers have studied the computational complexity of using LMSR to support various combinatorial betting languages. It is computationally hard to price most Boolean and permutation betting languages.[36] But LMSR can be operated efficiently for tournament betting[37] and taxonomy betting.[38]

3.2 Dynamic pari-mutuel markets

A dynamic pari-mutuel market (DPM)[39] is a dynamic-cost variant of a pari-mutuel market. It is a hybrid of a pari-mutuel market and a CDA, designed to provide infinite liquidity as a pari-mutuel market and to incentivize participants to reveal their information earlier, as in a CDA.

There are n contracts offered in a DPM, each corresponding to an outcome of v. As in a pari-mutuel market, traders who wager on the true outcome split the total pool of money at the end of the market. However, the price of a single share varies dynamically according to a price function. Thus, traders may want to purchase shares early to get a lower price. They can also sell their shares prior to the determination of the outcome for profits or losses. From a trader's perspective, DPM acts as a market maker in a similar way as cost-function-based market makers. A natural way for the DPM market maker to set contract prices is to equate the ratio of prices of any two contracts by the ratio of the number of shares outstanding for the two contracts. Then the cost function of DPM is:

$$C(\mathbf{q}) = \kappa\sqrt{\sum_{j=1}^{n} q_j^2} \tag{4.10}$$

while the instantaneous price function is:

$$p_i(\mathbf{q}) = \frac{\kappa q_i}{\sqrt{\sum_{j=1}^{n} q_j^2}}, \tag{4.11}$$

where κ is a positive free parameter. Unlike cost-function-based market makers, where payoff of a contract is fixed at \$1 if the corresponding outcome becomes true, payoff of a contract in DPM cannot be determined until the close of the market. If outcome k happens, the payoff per share for contract k is:

$$o_k = \frac{\kappa\sqrt{\sum_{j=1}^{n}(q_j^f)^2}}{q_k^f}, \tag{4.12}$$

where q_f is the quantity vector at the end of the market. If a trader wagers on the correct outcome, the price he pays is always less than κ and the payoff he gets is always greater than κ. It is natural to set $\kappa = 1$.

Unlike MSR and cost function-based market makers, where the market price of a contract represents the market probability of the corresponding outcome, instantaneous prices in DPM do not represent a probability distribution. The market probability of outcome i in DPM with the above-described cost, price and payoff functions is given by:

$$\pi_i(\mathbf{q}) = \frac{p_i(\mathbf{q})}{C(\mathbf{q})/q_i} = \frac{q_i^2}{\sum_{j=1}^{n} q_j^2}. \tag{4.13}$$

For traders whose probabilities are the same as the market probabilities, they cannot expect to profit from buying or selling contracts if the DPM market liquidates in the current state.

DPM needs the market maker to seed the market with some initial shares (money), which can be arbitrarily small, because the price function is not defined at $\mathbf{q} = 0$. Tech Buzz Game[40] was based on the DPM mechanism.

4 Conclusion

We have discussed several existing market mechanisms that can be used to operate prediction markets. They are: (1) call markets; (2) continuous double auctions; (3) generalized call mechanisms; (4) pari-mutuel markets; (5) market scoring rules; (6) cost-function-based market makers; and (7) dynamic pari-mutuel markets. The first three mechanisms are auctioneer mechanisms, while the last three are automated market maker mechanisms. In terms of the desired properties for information aggregation, pari-mutuel markets and all automated market maker mechanisms can provide infinite liquidity, but the auctioneer mechanisms more or less suffer from the thin market problem. Market scoring rules, cost-function-based market makers, and generalized call mechanisms can be used to operate combinatorial prediction markets to increase expressiveness in getting and processing information. However, for many combinatorial betting languages, the auctioneer's order matching problem and the market maker's pricing problem become computationally intractable. In auctioneer mechanisms and pari-mutuel markets, the market institution does not incur any loss. The automated market makers subsidize the market but have bounded worst-case loss.

Notes

1 Cason and Friedman (1997); Friedman (1993).
2 Satterthwaite and Williams (1993).
3 For illustration purposes, we assume that each order demands or supplies one share of the contract. The mechanism works for any limit orders.
4 Friedman (1993)
5 http://www.biz.uiowa.edu/iem.
6 http://www.intrade.com.
7 http://us.newsfutures.com.
8 Chen and Plott (2002); Chen *et al.* (2003).
9 Cowgill *et al.* (2008).
10 Bossaerts *et al.* (2002).
11 Lange and Economides (2007); Baron and Lange (2005).
12 Execution risk refers to the situation where desired trades either will not take place or will happen, but at unexpected prices, potentially affecting the optimality of the resulting portfolio.
13 Chen *et al.* (2007b).
14 Peters *et al.* (2007).
15 Fortnow *et al.* (2005).
16 Chen *et al.* (2007a).
17 Ghodsi *et al.* (2008).
18 Agrawal *et al.* (2008).
19 Chen *et al.* (2007a).
20 Rosett (1965); Weitzman (1965); Ali (1977); Snyder (1978); Thaler and Zhembou (1988).
21 Milgrom and Stokey (1982).
22 McCarthy (1956); Hendrickson and Buehler (1971); Savage (1971); Gneiting and Raftery (2007).
23 Brier (1950); Good (1952); Winkler (1967); Winkler (1969); Savage (1971); Winkler *et al.* (1996).
24 Nelson and Bessler (1989).
25 Murphy and Winkler (1984); Spiegelhalter (1986); O'Carroll (1977).
26 Matheson and Winkler (1976).
27 Savage (1971); Cervera and Munoz (1996); Gneiting and Raftery (2007); Lambert *et al.* (2008b).
28 Kilgour and Gerchak (2004); Lambert *et al.* (2008a).
29 Hanson (2003); Hanson (2007).
30 Chen and Pennock (2007).
31 Chen and Vaughan (2010).
32 Chen and Vaughan (2010).
33 http://inklingmarkets.com.
34 Chen and Pennock (2007).
35 Agrawal *et al.* (2009).
36 Chen *et al.* (2008a).
37 Chen *et al.* (2008c).
38 Guo and Pennock (2009).
39 Pennock (2004); Mangold *et al.* (2005); Chen *et al.* (2008b).
40 Mangold *et al.* (2005); Chen *et al.* (2008b).

References

Agrawal, Shipra, Zizhuo Wang and Yinyu Ye. Parimutuel betting on permutations. In *WINE '08: Proceedings of the 4th International Workshop on Internet and Network Economics*, 126–137, Berlin, 2008.

Agrawal, Shipra, Erick Delage, Mark Peters, Zizhuo Wang and Yinyu Ye. A united framework for dynamic pari-mutuel information market design. In *EC '09: Proceedings of the 10th ACM Conference on Electronic Commerce*, 255–264, New York, 2009.

Ali, Mukhtar M. Probability and utility estimates for racetrack bettors. *Journal of Political Economy*, 85(4): 803–816, 1977.

Baron, Ken and Jeffrey Lange. *Parimutuel Applications in Finance: New Markets for New Risks*. Palgrave Macmillan, 2005.

Bossaerts, Peter, Leslie Fine, and John O. Ledyard. Inducing liquidity in thin financial markets through combined-value trading mechanisms. *European Economic Review*, 46(9):1671–1695, 2002.

Brier, Glenn W. Verification of forecasts expressed in terms of probability. *Monthly Weather Review*, 78(1):1–3, 1950.

Cason, Timothy N. and Daniel Friedman. Price formation in single call markets. *Econometrica*, 65(2):311–345, 1997.

Cervera, J. and J. Munoz. Proper scoring rules for fractiles. *Bayesian Statistics*, 5:513–519, 1996.

Chen, Kay-Yut and Charles R. Plott. Information aggregation mechanisms: concept, design and implementation for a sales forecasting problem. Working paper No. 1131, California Institute of Technology, Division of the Humanities and Social Sciences, 2002.

Chen, Kay-Yut, Leslie R. Fine and Bernardo A. Huberman. Predicting the future. *Information System Frontier*, 5(1):47–61, 2003.

Chen, Yiling and David M. Pennock. A utility framework for bounded-loss market makers. In *UAI '07: Proceedings of the 23rd Conference on Uncertainty in Artificial Intelligence*, 49–56, 2007.

Chen, Yiling, Lance Fortnow, Evdokia Nikolova and David M. Pennock. Betting on permutations. In *EC '07: Proceedings of the 8th ACM Conference on Electronic Commerce*, 326–335, New York, 2007a.

Chen, Yiling, Lance Fortnow, Evdokia Nikolova, and David M. Pennock. Combinatorial betting. *SIGecom Exchange*, 7(1):61–64, 2007b.

Chen, Yiling, Lance Fortnow, Nicolas Lambert, David M. Pennock and Jennifer Wortman. Complexity of combinatorial market makers. In *EC '08: Proceedings of the 9th ACM Conference on Electronic Commerce*, 190–199, New York, 2008a.

Chen, Yiling, David M. Pennock and Tejaswi Kasturi. An empirical study of dynamic pari-mutuel markets: evidence from the Tech Buzz Game. In *WebKDD '08: Proceedings of the Workshop on Web Mining and Web Usage Analysis*, 2008b.

Chen, Yiling, Sharad Goel and David M. Pennock. Pricing combinatorial markets for tournaments. In *STOC '08: Proceedings of the 40th Annual ACM Symposium on Theory of Computing*, 305–314, New York, 2008c.

Chen, Yiling and Jennifer Wortman Vaughan. A new understanding of prediction markets via no-regret learning. In *EC '10: Proceedings of the 11th ACM Conference on Electronic Commerce*, 189–198, New York, 2010.

Cowgill, Bo, Justin Wolfers and Eric Zitzewitz. Using prediction markets to track information flows: evidence from Google. Working paper, 2008.

Fortnow, Lance, Joe Kilian, David M. Pennock and Michael P. Wellman. Betting Boolean-style: a framework for trading in securities based on logical formulas. *Decision Support Systems*, 39(1):87–104, 2005.

Friedman, Daniel. The double auction market institution: a survey. In Daniel Friedman and John Rust, eds., *The Double Auction Market: Institutions, Theories and Evidence*. Perseus Publishing, 1993.

Ghodsi, Mohammad, Hamid Mahini, Vahab S. Mirrokni and Morteza Zadimoghaddam. Permutation betting markets: singleton betting with extra information. In *EC '08: Proceedings of the 9th ACM Conference on Electronic Commerce*, 180–189, New York, 2008.

Gneiting, Tilmann and Adrian E. Raftery. Strictly proper scoring rules, prediction, and estimation. *Journal of the American Statistical Association*, 102(477):359–378, 2007.

Good, Irving J. Rational decisions. *Journal of the Royal Statistical Society, Series B (Methodological)*, 14(1):107–114, 1952.

Guo, Mingyu and David M. Pennock. Combinatorial prediction markets for event hierarchies. In *AAMAS '09: Proceedings of the 8th International Conference on Autonomous Agents and Multiagent Systems*, 201–208, Richland, SC, 2009.

Hanson, Robin D. Combinatorial information market design. *Information Systems Frontiers*, 5(1):107–119, 2003.

Hanson, Robin D. Logarithmic market scoring rules for modular combinatorial information aggregation. *Journal of Prediction Markets*, 1(1):1–15, 2007.

Hendrickson, Arlo D. and Robert J. Buehler. Proper scores for probability forecasters. *Annals of Mathematical Statistics*, 42(6):1916–1921, 1971.

Kilgour, D. Marc and Yigal Gerchak. Elicitation of probabilities using competitive scoring rules. *Decision Analysis*, 1(2):108–113, 2004.

Lambert, Nicolas, John Langford, Jennifer Wortman, Yiling Chen, Daniel M. Reeves, Yoav Shoham and David M. Pennock. Self-financed wagering mechanisms for forecasting. In *EC '08: Proceedings of the 9th ACM Conference on Electronic Commerce*, 170–179, New York, 2008a.

Lambert, Nicolas, David M. Pennock and Yoav Shoham. Eliciting properties of probability distributions. In *EC '08: Proceedings of the 9th ACM Conference on Electronic Commerce*, 129–138, New York, 2008b.

Lange, Jeffrey and Nicholas Economides. A parimutuel market microstructure for contingent claims trading. NYU Stern School of Business Discussion Paper No. EC-01-13, 2007.

McCarthy, John. Measures of the value of information. *PNAS: Proceedings of the National Academy of Sciences of the United States of America*, 42(9):654–655, 1956.

Mangold, Bernard, Mike Dooley, Rael Dornfest, Gary W. Flake, Havi Hoffman, Tejaswi Kasturi and David M. Pennock. The tech buzz game. *IEEE Computer*, 38(7):94–97, 2005.

Matheson, James E. and Robert L. Winkler. Scoring rules for continuous probability distributions. *Management Science*, 22(10):1087–1096, 1976.

Milgrom, Paul and Nancy L. Stokey. Information, trade and common knowledge. *Journal of Economic Theory*, 26(1):17–27, 1982.

Murphy, Allan H. and Robert L. Winkler. Probability forecasting in meteorology. *Journal of the American Statistical Association*, 79(387):489–500, 1984.

Nelson, Robert G. and David A. Bessler. Subjective probabilities and scoring rules: experimental evidence. *Journal of Agricultural Economics*, 71(2):363–369, 1989.

O'Carroll, F.M. Subjective probabilities and short-term economic forecast: an empirical investigation. *Applied Statistics*, 26(3):269–278, 1977.

Pennock, David M. A dynamic pari-mutuel market for hedging, wagering, and information aggregation. In *Proceedings of the 5th ACM Conference on Electronic Commerce*, New York, 2004.

Peters, Mark, Anthony Man-Cho So and Yinyu Ye. Pari-mutuel markets: mechanisms and performance. In *WINE '07: Proceedings of the 3rd International Workshop on Internet and Network Economics*, 82–95, 2007.

Rosett, Richard N. Gambling and rationality. *Journal of Political Economy*, 73(6):595–607, 1965.

Satterthwaite, M.A. and S.R. Williams. The Bayesian theory of the k-double auction. In Daniel Friedman and John Rust, eds., *The Double Auction Market: Institutions, Theories and Evidence*. Perseus Publishing, 1993.

Savage, Leonard J. Elicitation of personal probabilities and expectations. *Journal of the American Statistical Association*, 66(336):783–801, 1971.

Snyder, Wayne W. Horse racing: testing the efficient markets model. *Journal of Finance*, 33(4):1109–1118, 1978.

Spiegelhalter, D.J. Probabilistic prediction in patient management and clinical trials. *Statistics in Medicine*, 5(5):421–433, 1986.

Thaler, Richard H. and William T. Ziemba. Anomalies: parimutuel betting markets – race-tracks and lotteries. *Journal of Economic Perspectives*, 2(2):161–174, 1988.

Weitzman, Martin. Utility analysis and group behavior: an empirical study. *Journal of Political Economy*, 73(1):18–26, 1965.

Winkler, Robert L. The quantification of judgment: some methodological suggestions. *Journal of the American Statistical Association*, 62(320):1105–1120, 1967.

Winkler, Robert L. Scoring rules and the evaluation of probability assessors. *Journal of the American Statistical Association*, 64(327):1073–1078, 1969.

Winkler, Robert L., Javier Muoz, Jos Cervera, Jos Bernardo, Gail Blattenberger, Joseph Kadane, Dennis Lindley, Allan Murphy, Robert Oliver and David Ros-Insua. Scoring rules and the evaluation of probabilities. *TEST: An Official Journal of the Spanish Society of Statistics and Operations Research*, 5(1):1–60, 1996.

5 Information markets for decision-making

Performance and feasibility

*Karen Croxson**

1 Introduction

The success of any organization hinges on its ability to make wise decisions in an uncertain environment. When, where and how should scarce resources be committed to achieve institutional objectives? A product manager must devise a strategy commensurate with the profit objective. Politicians must decide how best to allocate the public purse to achieve policy goals. A disaster relief agency must make tough decisions as to where to concentrate efforts and supplies on the ground. Making the right decision requires weighing the effectiveness of alternative actions based on relevant information, which often is widely dispersed. It can be challenging to identify genuine insights and somehow aggregate these into a meaningful intelligence. Until recently, decision-makers have tended to rely on a number of traditional approaches to informing their decisions, including consulting experts, appointing groups to deliberate and conducting polls and surveys. We can think of each as an information mechanism – with its merits, but also recognized shortcomings. For instance, 'groupthink' tends to bias the views of committees; those we approach as 'experts' may not be true experts (and may not give honest opinions); a simple poll fails to weight individual opinions according to how informative they actually are.

In recent years a band of innovative organizations, including Hewlett Packard, Google, General Electric and Microsoft, has experimented with a new mechanism – the information market. Information markets essentially are in-house betting markets. They are designed specifically to tap the knowledge of a dispersed 'crowd' and aggregate this into an accurate forecast. Participants in an information market buy and sell assets whose payoffs are tied to the realization of a future event, perhaps the effectiveness of a current sales campaign or even the success of a merger. The market price is interpreted as a collective prediction. There are sound theoretical reasons to expect this forecast to be more reliable than forecasts provided by alternative methods, by getting participants to 'put their money where their mouth is' without fear of recrimination (thanks to anonymous trading), a market addresses many of the difficult incentive problems that can undermine more traditional information mechanisms. Public betting markets have a rich tradition in forecasting political elections and sports events, performing impressively against polls and pundits (Berg *et al.*, 2008; Vaughan Williams, 2005); many laboratory experiments have demonstrated the ability of markets to aggregate the information of traders effectively (Plott and Sunder, 1982, 1988; Forsythe *et al.*, 1982; O'Brien and Srivastava, 1991), and where information markets have been applied inside corporations they appear to have delivered some encouraging results (Ortner, 1998; Plott and Chen, 2002; Cowgill *et al.*, 2009).

This combination of theoretical appeal and promising empirical evidence has led to speculation that in-house information markets will revolutionize decision-making across private and public sector organizations.[1] Meanwhile, the business columns of international magazines and newspapers have led many high-profile discussions of the potential for the 'corporate prediction market' to become a killer Web 2.0 application.[2] Despite all the enthusiasm and expectation, real-world utilization of information markets remains far behind the hype. The vast majority of the business world is showing little interest in the internal market as a serious alternative (or even complement) to standard forecasting methods, even for more pedestrian business matters. Moreover, even among the handful of firms embracing the technology most enthusiastically, none is known to use the mechanism to guide major strategic decisions. The public sector, meanwhile, seems barely to have begun to experiment with the concept, despite the potentially huge wins from 'policy markets'.[3] This chapter reviews some of the evidence surrounding the comparative performance and feasibility of information markets, discusses some of the practical challenges inhibiting their mainstream deployment inside organizations, and touches on priorities for future work.

2 Traditional information mechanisms

2.1 Expert opinion

A time-honoured approach to informing decisions is to seek out an expert opinion. Appealing to expertise has intuitive appeal; most of us would prefer to act on the medical advice of a qualified doctor than rely on the judgement of our friends or neighbours. There are difficulties with the expert model, however. Whereas in some cases it may be trivial to identify the right person – their background or specialist training might mark them out for the role – in many other cases identifying a true expert is challenging, and those who appear to be expert in a matter may possess little actual expertise. Tetlock (2005) describes a 20-year study in which several hundred experts drawn from many fields, from professors to journalists, were asked to make thousands of predictions about the future. The predictions turned out to be only marginally more accurate than chance, and the most recognized experts – those who advise governments and business leaders, appear on television and are regularly quoted in newspapers – performed particularly poorly. Even when a real expert has been located, simply asking for an opinion may not be the best way to ensure an honest response. Experts can suffer conflicts of interest: they may desire to tell us what they think we'd like to hear (Prendergast, 1993), or feel it is safer to herd with other experts (to avoid being wrong when everyone else is right).[4] A narrow focus on conventional circles of expertise might cause valuable knowledge to be missed; recent work has shown that official forecasts of unemployment can be improved using insights from internet search data (D'Amuri and Marcucci, 2009). Similar results have emerged for the detection of influenza (Ginsberg *et al.*, 2009; Polgreen *et al.*, 2007). Instead of relying on a single, potentially biased, opinion we might think about consulting several experts and somehow pooling their opinions. This may go some way towards mitigating biases, although it still leaves open the problem of distinguishing true experts. It also raises the non-trivial question of which pooling mechanism should be used.[5]

2.2 Group deliberation

Appointing a group to deliberate would appear to provide an attractive basis for decision-making: gather together individuals with relevant perspectives, encourage them to engage with each other, share insights, resolve differences, and so converge on an enlightened collective viewpoint. Certainly, meetings are used heavily inside most organizations. But in practice, important limitations can arise, particularly where meetings are face-to-face. Minorities with relevant information tend to be silenced too readily by social pressure, and junior members of a meeting can be reluctant to say what they really think. A troubling 'groupthink' can emerge,[6] a lack of anonymity can inhibit honesty in communication, and it is often necessary to exclude some who may have relevant information – a many-way discussion can become quite unmanageable as the number of people involved grows large.

2.3 Polls and surveys

For some issues, it might be useful to conduct a poll across the organization. Polls can provide a quick snapshot of broad opinion, which could be informative, particularly if the poll is anonymous. Instant polls are becoming an everyday part of life in some organizations, spurred by the development of handheld voting devices. These devices were first used in the 1960s by motion picture and television studios as a way to gauge audience responses to unreleased films, TV shows and commercials. Nowadays, they are used inside companies to take the pulse of opinion during meetings.[7] A theoretical drawback with polls is that, by default, they weigh individual opinion equally, and yet people may be quite differently informed. A potential response would be to assign different weights to individual opinions to reflect their relative expertise, but in this case the problem of identifying experts resurfaces. Polls also fail to reward people when their opinion is accurate and so may inhibit participation and truthful revelation of information. Surveys can be used to give a more nuanced view of opinions, but they tend to suffer from similar weaknesses: they fail to reward individuals for reporting their insights truthfully, and they still leave the problem of how to interpret and weight the responses. Surveys also need to be designed, and their implementation can be very costly.

3 The promise of information markets

Information markets are betting markets established specifically to aggregate dispersed information into a collective forecast. They go by many other names, including prediction markets, event futures, event derivatives and virtual stock exchanges. Participants in an information market are traders – they buy and sell contracts which have payoffs tied to the realization of a future event. Suppose a computer games executive wishes to know whether a new game will ship on time. She could set up an internal information market to help her estimate whether things are on track. After she has procured some simple software to run her market (which is available off-the-shelf from many specialist technology providers, or might be built in-house), she defines the contract 'Product X will ship on 1 November 2011 at the latest', she stipulates that this contract will pay out $1 (which could be virtual currency) if the designated success occurs and $0 otherwise. Their next step is to invite a group of participants to trade the contracts, perhaps providing each with an initial endowment. Employees from around the firm might be invited to trade, or only those closely acquainted with Product X. To kick things off, the executive might set the initial contract price at $0.50, implying a 50:50 chance that the product ships on time. Some of the participants may find this belief too pessimistic based on what they know and so will be incentivized to buy contracts in the market. Others may feel even more pessimistic about delivery than the current price implies and so will wish to sell contracts. As individuals trade in the market the contract price will move around to reflect their information. After a few weeks the contract might be priced at $0.80, from which the product manager should infer that the chances that Product X ships on time are around 80%.

There are good reasons to think that a forecast generated in this way could be highly accurate - potentially more reliable than a forecast generated in any other way. Economists have long recognized that a byproduct of markets designed for speculation and hedging is that they can aggregate private information effectively (Hayek, 1945):[8]

- Markets enable insights to be gathered from a diverse crowd of participants, and diversity has been shown to be important for problem solving (Surowiecki, 2004; Page, 2007).
- Traders can submit their views independently and anonymously in a market – unlike in deliberative group settings.
- Markets provide the right incentives to participants to reveal their information quickly and truthfully – those who have information and are first to trade on this can profit.

- Markets provide incentives for research and information discovery – individuals who actively acquire information can earn additional profits through trading.
- The market provides an algorithm for aggregating opinions into a single collective viewpoint.

Markets offer predictions that update 'in-running'. They have been described as a 'pull' forecasting mechanism, as compared to the 'push' mechanisms offered by polls and surveys.[9] As new information emerges, traders in prediction markets have the incentive and opportunity to react quickly by changing their positions, and so market price should move rapidly to incorporate any news.

Recent years have seen the creation of many public prediction markets, dealing with election results through to the winners of reality-TV contests. Forecast performance has been impressive, with market predictions routinely topping those of professional forecasters and polls in head-to-head contests. Perhaps the best known information markets are the Iowa Political Markets, created by academics at the University of Iowa in the late 1980s. They allow the public to bet on political outcomes such as US presidential elections. The predictions from Iowa markets have beaten opinion polls and political pundits remarkably consistently over the years (Forsythe *et al.*, 1992; Berg *et al.*, 2008). Following this success, Iowa researchers have developed markets to forecast outbreaks of avian flu and to predict the Federal Reserve's monetary policy. Elsewhere, Hollywood markets invite the public to predict opening weekend box office sales and pick Oscar winners – anyone can go to the Hollywood Stock Exchange website (www.hsx.com), sign up for free, and automatically collect an endowment of Hollywood dollars to buy and sell films and actors. Around 1.8 million people are now registered to trade and the predictions from these markets tend to be more accurate than those of film critics, even though only play-money is at stake (Pennock *et al.*, 2001; Spann and Skiera, 2003). Meanwhile, real-money prediction markets have existed for sporting events for some time. The online betting exchange Betfair (www.betfair.com) is the world's largest prediction exchange. It is reported to have over two million members and offers real-money markets related to thousands of sports events (and many other events of popular interest). Studies to date have reported encouraging evidence regarding the efficiency of the exchange's prices.[10]

The earliest example of corporate information markets are thought to be the markets academic economist Robin Hanson helped establish at technology provider Xanadu in 1990. One of the claims traded related to the delivery date of the firm's product: 'Xanadu will deliver its product before Premier Deng of China dies.' Deng died before the product was delivered.[11] Since then, many more businesses have begun to experiment with their use internally. The markets seem to be used primarily to forecast such matters as whether a project deadline will be met, whether a sales target will be exceeded, or what a competitor will do. Typically, a relatively small group of employees is invited to trade, and often play-money and prizes are the only incentives offered. The results reported so far are encouraging. An early pioneer in this area was Hewlett Packard. It began experimenting with internal markets in the late 1990s, leveraging its crowd of employees to forecast printer sales. The forecasts from HP's information markets beat the forecasts provided by its sales department – the 'experts' (Plott and Chen, 2002). Google seems to have conducted the largest experiment with corporate prediction markets to date, launching its first markets in 2005 and going on to deploy thousands internally. Its markets are used to forecast product launch dates and a range of other outcomes. Cowgill *et al.* (2009) analyse data from Google's markets for the period 2005–2007. Though they find evidence of biases on the part of traders (such as optimism bias, which seems strongest in newer recruits), they report that markets are reasonably efficient and become less biased over the study period as collective trading experience increases. Microsoft has explored prediction markets internally, as have retailer Best Buy and several other leading firms, including major pharmaceutical players Pfizer, Novartis, GSK and Eli Lilly.

4 Barriers to utilization

Surprised by the unfulfilled potential of information markets, James Surowiecki wrote in his 2004 best-seller, *the wisdom of crowds*, that: '...the most mystifying thing about markets is how little interest corporate America has shown in them ... companies have remained, for the most part, indifferent to this source of potentially excellent information (Surowiecki, 2004: 21–22).' By popularizing the notion of collective wisdom, Surowiecki himself did much to acquaint real-world decision-makers with the idea of information markets, leading to a leap in awareness among executives. In the six years since then, numerous articles in leading business magazines and the popular press have addressed the promise of internal markets, ensuring that their public profile remains high. Despite all this, there remains little evidence to suggest that in-house markets are becoming mainstream. Over the past few years, consultancy McKinsey has surveyed executives from a broad range of industries, regions and functional specialties about their usage of Web 2.0 applications such as blogs, podcasts, social networks and prediction markets. In the 2009 survey, only a small minority of the approximately 1,700 respondents reported that prediction markets were deployed inside their organizations – 9 per cent of executives, largely unchanged from the previous year (McKinsey, 2009).[12] Just under half of these described their institution's use of markets as 'evaluating or running limited trials' (as opposed to 'using it in our business') and around half were only, at most, somewhat satisfied with the technology.[13] Focusing on government applications, the evidence is even less encouraging. There are few signs of experimentation with prediction markets within public sector organizations, let alone serious adoption.[14] In comparison to the media hype surrounding their potential to revolutionize decision-making, it would appear that prediction markets are experiencing a peculiar 'failure to launch'.

Will we ever see organizations leverage markets to guide their large-scale strategic bets, such as whether to merge, sack the current CEO or push into a new geographic territory? Will their deployment for more pedestrian purposes ever become routine? In the next section, we discuss some of the design challenges surrounding the application of internal markets and other practical impediments to their proliferation inside organizations. We consider where and how some of these issues might be mitigated.

4.1 Restrictions on applicability

Information markets cannot be deployed to address every issue of interest. Compared to more traditional information mechanisms, more onerous conditions must be satisfied for a situation to be amenable to the use of markets. Specifically, it must be feasible to:

- define an uncertain event unambiguously;
- write contracts related to a (small) number of mutually exclusive and mutual exhaustive possible outcomes;
- resolve uncertainty by a specified point in time (ideally not too far into the future as long-term markets tend to be less motivating);
- settle the market against objective criteria.

This means that rather than listing a contract with the wording 'Weapons of mass destruction are not in Iraq', which specifies no end point at which the bet can be settled, contracts of the form 'WMD will have been found by date Z' should be used. Nor can the definition of the event change once the market is in progress: Ortner (1998) describes an internal prediction market used to forecast whether a software project would be delivered to the client on schedule. At some stage, the client changed the deadline, creating problems for the operation of the market. Markets will only work well only when there is reasonable clarity about and confidence in the contract design. Clearly, not all questions of interest will be compatible with the above conditions, whereas more traditional information mechanisms can be applied to less well-defined issues, including matters of opinion. Even so, a vast range of organizational issues will be amenable to markets, including potentially many 'big-ticket' questions such as: 'Will medical device Y be recalled by the end of this quarter?', 'Will our rivals GHI Inc. and JKL Inc. merge in the next six months?' or 'Will our health reform lower hospital admittances by Z% this year?'.

4.2 Limited empirical evidence

At least some of the reluctance to implement information markets can be explained by the lack of clear-cut evidence on their performance against alternatives. A related issue is confusion and uncertainty related to design choices. The number of academic articles on the topic of information markets has grown rapidly in recent years,[15] but there remains limited clarity surrounding what should be expected from markets in specific real-world settings and how exactly markets should be configured.[16]

A number of studies have confirmed the ability of information markets to aggregate information effectively in the laboratory (Plott and Sunder, 1982, 1988; Forsythe *et al.*, 1982; O'Brien and Srivastava, 1991), but other experiments have demonstrated the existence of problems such as bubbles and false equilibria (Smith *et al.*, 1988; Camerer and Weigelt, 1991; Noeth *et al.*, 1999; Lei *et al.*, 2001; Hussam *et al.*, 2008). While laboratory settings enable the experimenter to assign private information to individual participants, allowing the performance of information mechanisms to be tested cleanly, they come with an obvious downside: participants are typically inexperienced students confronted by abstract problems. As such, it is unclear whether the laboratory performance of information mechanisms extrapolates to corporate or policy contexts.[17]

Other researchers have investigated performance in the field (Forsythe *et al.*, 1992; Chen *et al.*, 2005; Christiansen, 2007; Cowgill *et al.*, 2009), and some of these have compared the accuracy of market forecasts to those generated by other information mechanisms. For instance, Chen *et al.*, (2005) analyse predictions from two public information markets for NFL games played in 2003, comparing these to 'expert' opinion pools for the same events. The opinion pools are constructed using linear and logarithmic aggregation functions to combine the subjective probability judgements of 2,000 self-identified experts. The authors find that the predictions from information markets are as accurate as the pooled expert assessments for the same time-point ahead of the game. But there are limitations to studies that compare mechanisms run side-by-side. When the mechanisms studied estimate the same outcomes (and may potentially even share some participants), this undermines the potential for clean inference regarding the efficacy of either mechanism. Information may leak from one mechanism to the other in a way unobserved by the researcher. For greater clarity, future work should pursue more systematic comparison of mechanisms in real-world settings. Graefe (2009) provides a useful recent contribution. In a field experiment the author compares the Delphi method of structured group deliberation with prediction markets, taking care to ensure that none of the market participants concurrently took part in the Delphi mechanism. Prediction markets are found to work as well as the Delphi method.

Difficult design choices have to be made in setting up an information market, but despite papers such as Spann and Skiera (2003), still little is understood about the impact of market design features on information accuracy. Two important open questions relate to participation and incentives.[18]

Participation

If active traders are few and far between, the market may become too thin to yield accurate predictions. However, recent research has suggested that, depending on the market mechanism implemented, the lower bound on numbers may be quite low (Christiansen, 2007; Chen *et al.*, 2001).

It might seem intuitive that participation should be restricted to those with some expertise about the subject matter. Assuming experts can be reliably identified, including non-experts in many conventional information mechanisms, such as meetings, would seem to add little obvious value. Prediction markets potentially are very different in this respect. The presence of uninformed traders (referred to as 'noise traders' in the finance literature) may be necessary for the market to be viable (Wolfers and Zitzewitz, 2006) and may actually increase the incentives for those with information to bring this to the market since the presence of uninformed others implies an opportunity to trade profitably on knowledge.[19] At the same time, the jury is still out on just how much non-informed trading is desirable: while some empirical studies have lent support to the view that securities mispricing is greater in illiquid markets (Kumar and Lee, 2006; Sadka and Scherbina, 2007; Chordia *et al.*, 2008), in the theoretical work of De Long *et al.* (1990), liquidity is a proxy for non-informational trading (noise trading), which may harm informational efficiency, and some recent empirical analysis of public prediction markets appears to support the idea that greater liquidity can worsen mispricing (Tetlock, 2008; Hartzmark and Solomon, 2010). Clearly, further research is needed in this area, with more investigation of the impact of non-informed trading in internal applications. Would using individual market trading performance as an indicator of expertise and then pooling 'expert' opinions yield a more informative mechanism? Work by Chen *et al.* (2001) suggests that it might.

In practice, those setting up prediction markets have taken a mix of approaches to the question of who to include. When creating the Iowa Election Markets in 1988, academics at Iowa University opted to open these markets to the general public. However, when creating their first Iowa Health Market to predict seasonal influenza (in 2004), it was decided to restrict participation to registered members of the medical community. The University of Iowa now runs two types of market: public (play money) and private (real money). Applications to play its private markets, which include many health markets, are reviewed and approved by market managers on a case-by-case basis, with those viewed as the best possible traders selected to play.[20] When retailer Best Buy experimented with internal information markets for forecasting business outcomes such as sales, the company was careful to include a wide base of participants with diverse operational knowledge, and not simply in-house forecasting experts. This approach respects the finding that sufficient cognitive diversity is important for a crowd to be wise (Page, 2007).[21]

Incentives

One of the biggest potential advantages associated with prediction markets is that participants are forced to 'put their money where their mouth is'. Compared to other information mechanisms, this should reduce biased messages. Establishing a real-money information market raises legal, financial and ethical issues, however; in many jurisdictions, including the United States, gambling is heavily restricted.[22] A group of prominent academics has called publicly for regulators to make provisions for real-money prediction markets on the grounds of social interest (Arrow *et al.*, 2007), but for now most companies experimenting with in-house information markets have chosen to proceed cautiously using virtual money and occasional prizes. Many economists would expect a play-money market to provide weaker incentives for information acquisition and truthful revelation since traders have less at stake. A large number of laboratory studies, beginning with Siegel (1961), have supported the idea that real money is required to ensure truth-revealing incentives in experiments. However, promising results from several play-money markets inside corporations have led some observers to suggest that participants derive sufficient motivation from the pursuit of 'winner status'.[23]

A few recent investigations explore the link between incentives and prediction market performance (Servan-Schreiber *et al.*, 2004; Rosenbloom and Notz, 2006; Luckner, 2007; Gruca *et al.*, 2008), but the evidence remains inconclusive. Servan-Schreiber *et al.* (2004) compare the predictions of TradeSports, a real-money market, and NewsFutures, a play-money market, finding that the play-money markets performed as well as the real-money markets. Rosenbloom and Notz (2006) also compared the predictions from TradeSports with those from NewsFutures; they report that there was little to separate the predictions for NHL games, but that the real-money market delivered more accurate predictions for non-sports events. The recent study by Gruca *et al.* (2008) investigates the impact of incentives on prediction for movie success. The authors compare predictions from the play-money market, Hollywood Stock Exchange, with those from the real-money Iowa Electronic Market for the same movies. Contrary to Rosenbloom and Notz (2006), they find no statistically significant difference between the accuracies of the two markets for these non-sports events. As noted previously, however, a limitation of studies which compare predictions from markets for the same events is the possibility for information leakage across mechanisms. This constrains the ability to deliver clean inference. Rosenbloom and Notz (2006) have suggested that a partial explanation for their results could be the existence of publicly available betting odds for the sports events they study; these odds could have helped traders in both the real-money and play-money markets. The impact of incentives in information markets deserves further research, ideally involving randomized trials inside organizations. In the context of enterprise information markets, it will be important also to consider how incentives compare to salaries, awards or other workplace incentives.

4.3 Sensitivity of market information

Perhaps the biggest barrier to the mainstream deployment of internal corporate markets relates to the sensitivity of the information generated. Managers may have concerns about the release of information regarding:

- the nature of the organization's problem (the fact that the firm is seeking product ideas, improved sales forecasts or entertaining the idea of a merger);
- the prevailing market forecast.

In their default implementation, prediction markets reveal to all participants an uncensored 'in-running' (and potentially very accurate) forecast related to the event of interest. However, this disclosure often sits uncomfortably with an institutional desire to bias official forecasts and potentially to conceal these from employees. In many interesting cases, it may simply be unacceptable to have an unbiased metric about sensitive organizational outcomes visible to regular employees this may impact organizational morale adversely or be leaked to the stock market in violation of insider trading provisions. A third worry is that sensitive intelligence might reach competitors.[24] We expand briefly below on the first two of these concerns.

Motivation and morale

Targets may be set to stretch employees and are hence potentially unrealistic by design. Evidence from academic studies lends support to this approach: psychologists have demonstrated the existence of the Pygmalion effect, which refers to the tendency for actual performance to converge to the positive expectation held by others (Eden, 1984).[25] Equally, morale may be sensitive to expectations about external market conditions or a firm's ability to retain its key staff. Consider an internal market set up to forecast employee retention which is predicting an exodus of staff over the coming quarter. A visible prediction of this nature could become a destabilizing focus within the organization, undermining the ability of leaders to manage.[26]

Complications related to 'insider trading'

Some managers have identified insider trading rules as an important factor discouraging them from using information markets to support 'big ticket' business decisions, such as whether to merge, introduce new products or expand into a new territory (Hanson, 2008). If employees receive key corporate information that is not publicly available, they become 'insiders' in a legal sense. If they were to trade on this information in the market for the company's stock, they would engage in 'insider trading' and civil or criminal sanctions might follow. Mat Fogarty, CEO of prediction markets provider Crowdcast, explains:

> The concern is that the information coming from the PM is so powerful that all recipients of that data are made insiders. Also, if management is aware of any 'material' information, it should disclose this to its shareholders. PMs are designed to deliver plenty of 'material' information.[27]

Bell (2009) has clarified the legal issues surrounding prediction markets and current insider trading provisions and suggests practical actions firms might take to mitigate problems associated with this potentially thorny issue. One idea is to create separate markets that could be traded only by officers of the company and other existing 'insiders', so that the company's existing framework of controls for inside information could be applied. Another suggestion is that the firm admonish participants in its private markets against trading on the information generated. Additionally, participants could be notified that claims and prices constitute the corporation's 'trade secrets', which, according to Bell, would give the firm a misappropriation claim against anyone who trades the company's shares based on its internal market information (rather than a joint liability for their insider trading). An interesting fourth suggestion is that firms might seed their internal information markets with a number of fake claims and prices, so that participants cannot tell which are real. Only traditional insiders would be told which claims were genuine and which were decoys. Looking beyond the current legal regime, Hanson (2008) has argued for changes to existing insider trading rules to allow firms to explore the gains from prediction markets more freely while still preserving the benefits traditionally associated with the existing rules that limit insider trading (the encouragement of investment in public corporations).

Reflecting the above concerns, some corporate adopters of information markets have shown interest in the possibility of preventing employees from viewing the current market price. Crowdcast, a commercial provider of prediction market solutions, has developed a system of 'blind betting', in which participants submit their predictions without observing the current market consensus.[28]

4.4 Fears about manipulation

A concern often raised in the context of discussions about information markets is the possibility of manipulation (Wolfers and Zitzewitz, 2004). The fear is that participants might be tempted to manipulate either the event outcome itself (in order to profit in the market) or market prices (in order to influence a real-world decision that is to be informed by the market price).

Manipulation of event outcomes

In some applications, the outcome of interest is exogenous to those trading the prediction market, as when members of the public bet on the outcome of a soccer match, or company employees speculate on the actions of a competitor firm. But in many interesting real-world implementations, those trading a prediction market will themselves have some influence over the outcome of interest. An obvious case arises where a small group of employees working on a particular project is invited to predict whether the project will be completed on time. The fact of their participation in the information market may create perverse incentives for employees to manipulate the outcome of the project. See Hanson (2006b) and Wolfers and Zitzewitz (2006) for informal discussions of outcome manipulation. A first formal analysis is provided by Ottaviani and

Sorensen (2006). They show how, in theory, outcomes might be manipulated by participants in corporate prediction markets.

Manipulation of market prices

An additional potential worry is that participants might feel tempted to manipulate prediction market prices, particularly where high-stakes decisions are to be based on these (Wolfers and Zitzewitz, 2006). This scenario has been analysed theoretically by Hanson and Oprea (2007) and Hahn and Tetlock (2007). Several empirical studies suggest manipulation is likely to be ineffective (Rhode and Strumpf, 2008; Hanson *et al.*, 2006). For instance, Rhode and Strumpf (2008) analyse an attempt to manipulate the price of a Kerry victory on the public market TradeSports in 2004, as well as their own attempts to manipulate prices on the Iowa Electronic Markets in 2000. These manipulation efforts seem to have had only a very short-lived effect on prices. It has even been suggested that manipulation may enhance market accuracy (Hanson and Oprea, 2007). Price manipulation remains an ongoing area of research with some recent investigations pointing to its potential effectiveness in particular settings (Veiga and Vorsatz, 2009, 2010).

In his book, *Predictocracy*, Michael Abramowicz (2007), offers the following advice on how to safeguard against market manipulation in practical settings:

> First, where there is a discrete group of potential manipulators, those individuals can be barred from participation. Of course, there is always a danger that these potential manipulators can pay off other market participants, but legal or contractual sanctions can reduce that possibility. Second, prediction markets might be limited to a group of authorized traders who are believed to have no incentive to manipulate the outcome.[29]

Others have suggested that incentives associated with the market should be kept below those associated with achieving the outcome of interest. That is, the personal return to ensuring that organizational goals are met should comfortably exceed any reward attainable from betting against these in the prediction market.

Some of those engaged with the practical deployment of information markets inside organizations appear to take the view that the fears surrounding manipulation are overblown, at least for play-money markets.[30]

4.5 Difficulties sustaining participation

Users often report difficulties with the trading interface. This needs to be kept as simple as possible. Some participants are simply uncomfortable with the trading metaphor, and the idea of translating their knowledge into a price (Green *et al.*, 2007). Adam Siegel at Inkling, another platform provider, has observed a shift away from trading screens that look like stock exchanges:

> When we first launched Inkling, we were greeted with great skepticism because our application looked nothing like a stock trading platform. Now if you look at the newcomers in the space, they all try to highlight 'ease of use' as their differentiator.[31]

Once participants have mastered how to play an information, the challenge becomes how to sustain their interest. One strategy is to offer random prizes for participation, as well as winner prizes. The utilization of information markets has so far been almost exclusively confined to forecasting near-term events, and it has been suggested that markets, at least in their current incarnation, are not well-suited to forecasting long-term outcomes because of the difficulties sustaining participation:

> Incentives lose power if the payoff is too remote, and feedback is important for driving participation and performance. Forecasting a result within a few quarters seems to work, but over a year begins to feel like a stretch. We are experimenting with alternative market structures that might help forecast the distant future while paying incentives more quickly.
>
> (Hopman, 2007)

4.6 Perceived lack of legitimacy

Finally, information markets may fail because of a lack of perceived legitimacy. Sunstein (2006b) notes that information markets remain relatively unfamiliar and suggests that their use might breed confusion and distrust. By contrast, deliberation has been found to increase confidence and decrease variance in the group's prediction, which gives deliberative groups greater perceived legitimacy. Recent work by Graefe (2009) has sought to compare the acceptability of different information mechanisms in laboratory experiments. Participants were observed to discount market results more than those from other forecasting mechanisms, despite the fact that this harmed accuracy. Some may have a simple lack of faith in market predictions, whereas others may feel positively threatened by their arrival: information markets throw into question the role of in-house experts, and the forecasts generated may temper executive decisions, thereby challenging existing hierarchical structures. For these reasons, management and employees may be hostile to the introduction of markets and these may fail to become integrated into the normal workflow. When this happens, time spent trading may be seen as a distraction from proper work duties, rather than a valuable exercise. In GE's experimental idea futures markets, limits were imposed on trading hours; to ensure that the market did not interfere with regular work time, participants were asked to trade before or after work, during lunch or for only a few minutes at a time during work hours (LaComb *et al.*, 2007).

Jeff Severts, who introduced prediction markets at US consumer electronics retailer Best-Buy, considers it vital to secure executive buy-in for such initiatives:

> support from senior executives is essential if you want to issue contracts on anything that might be controversial. 'Air cover' is a must or you'll find yourself trading on what kind of casserole we're having in the cafeteria on Thursday.
>
> (Dye, 2008)

One way to improve openness of management to information markets might be through the education of current and future executives. Business schools could play a greater role in raising awareness of the problem of aggregating dispersed information inside organizations and the merits and limitations of mechanisms to achieve this. Andrew McAfee of Harvard Business School has incorporated examples of prediction markets into some of his teaching to show executives how Web 2.0 can be used to tap collective intelligence.[32] Daphne Raban and Dorit Geifman have experimented with prediction markets in MBA teaching at the University of Haifa's School of Management, exposing students to the topic through web-based demonstrations. Their recent article (Raban and Geifman, 2009) discusses the pedagogical value of this approach.

5 Concluding remarks

Recent years have seen active experimentation with internal information markets, leading some to foresee that they will one day revolutionize decision-making. On current evidence, markets are some way off delivering on this expectation. Despite many years of experimentation on the part of innovative firms, with many encouraging results, information markets have so far failed to become established as a mainstream forecasting solution. Nor are there any signs that those who utilize markets currently are deploying these to guide 'big ticket' business decisions. If markets provide superior forecasts then the biggest gains will come from using them to inform key strategic choices, such as whether to merge with a rival or replace the CEO.

This chapter has highlighted some of the advantages of information markets over other mechanisms, but also the many practical barriers to their wider deployment in real-world organizations. Potentially, markets provide an ingenious solution to the incentive problems that can beset alternative approaches to informing organizational decision-making. At the same time, markets raise novel practical challenges; they are not suitable for all settings, and need to be designed and implemented carefully and sensitively to be effective. Markets may leak sensitive information in a way that other mechanisms do not (or do less), and this in turn can impact morale and motivation negatively, create legal complications by turning participants into 'insiders' and hand competitors vital commercial intelligence. Markets may be subject to manipulation (though the fear of this may be greater than the reality), and it can be challenging to sustain participation. However, executives and policy-makers must keep in mind the relevant counterfactual. The practical performance of markets should not be compared to some absolute ideal, rather to the merits and limitations of real-world alternatives. Holding meetings, relying on 'expert' opinion, conducting surveys, or polling employees are all information mechanisms with costs and potential weaknesses. Often, their limitations are not sufficiently appreciated and challenged.

To date, much of the experimentation with information markets has been carried out either inside laboratories or by corporations behind closed doors. Producing accurate forecasts is essential to the success of all organizations and more of these should consider collaborating with academics to road-test competing information mechanisms scientifically, ideally allowing randomized trials within their own realistic settings. Organizations should look to do this as part of a broader information-based strategy (Davenport, 2009).[33] Prejudices against information markets, related to negative feelings about 'gambling' in the workplace, the threat to existing authority, or simply the counterintuitiveness of crowd wisdom, could be challenged more aggressively through education. In particular, business schools could play a big part in stimulating practitioners to engage with the concept of information markets through classroom experience and open, scientific debate.

Notes

* This work was completed at New College, Oxford University, and the Oxford-Man Institute of Quantitative Finance. I am grateful to both institutions for their generous support. Parts of the chapter build on earlier background research carried out for a study on the performance of distributed problem-solving networks, which was supported by the Oxford Internet Institute and McKinsey & Company. Several colleagues provided valuable comments and input, particularly Bill Dutton and David Bray at the Oxford Internet Institute, Michael Chui and James Manyika at McKinsey, and Clare Leaver and Ian Jewitt at the Department of Economics at Oxford.

1 Sunstein (2006a) makes the case for a wide range of information markets to address a range of important issues, including predicting damage from natural disasters. Meanwhile, Robin Hanson has suggested that 'Decision markets will one day revolutionize governance, both public and private' (excerpt from Robin Hanson's review of Abramowicz (2007), full text available online at www.overcomingbias.com/2008/01/predictocracy.html).

2 See, for instance, coverage in *The Economist* (2005), the *New York Times* (Lohr, 2008), and the *Wall Street Journal* (Dvorak, 2008). The *Special Interest Group on Prediction Markets* at www.forecastingprinciples.com provides an updated overview of media coverage related to information markets.

3 The likely gains from markets for public policy have been articulated by several academic economists (Hanson, 2006a; Ledyard *et al.*, 2006; Hahn and Tetlock, 2005).

4 The tendency of experts to herd has been observed empirically in many settings, including economic forecasting: 'economic forecasters all tend to be wrong in the same way. Their incentives to flock together are obvious enough.' Tim Harford, *Financial Times*, 9 August, 2008. Models of 'career concerns' can explain this phenomenon theoretically. See, for example, Scharfstein and Stein (1990) for an agency theoretic model in which concern for professional reputation causes experts to mimic the decisions of other experts.

5 Opinion pools can be classified into two broad categories: mathematical approaches and behavioural approaches (Clemen and Winkler, 1999). In mathematical approaches, the opinions of individual experts are expressed as subjective probability

distributions over outcomes of an uncertain event. They are combined through various mathematical methods to form an aggregated probability distribution. The important assumption of behavioural approaches is that, through exchanging opinions or information, experts can eventually reach an equilibrium in which further interaction won't change their opinions. The Delphi technique (Linstone, 2007) is a well-known behavioural approach. Both mathematical and behavioural approaches have advantages but also drawbacks. See Chen *et al.* (2005) for further discussion.

6 Armstrong (2006) discusses many examples of behaviour in face-to-face meetings that can lead to biased results, including the tendency for personal influence over the group outcome to be related to how loudly a person speaks, tone of her voice and physical appearance. See Sunstein (2006) for further insights regarding deliberative groups. Deliberative groups amplify cognitive errors, and fall prey to informational cascades and group polarization. Structured group deliberation processes, such as the Delphi method, have been designed to try to mitigate some of the shortcomings of unstructured face-to-face meetings.

7 US real estate company, Forest City Enterprises Inc., uses instant polling to improve the decision-making in its communications department: 'When evaluating materials for the company newsletter, website or other communications tools, the company uses the anonymity of the technology to gain input from the entire team.' http://marketingpr.suite101.com/article.cfm/audience_response_systems_for_employee_opinion.

8 The following discussion draws on points made by Wolfers and Zitzewitz (2004).

9 http://future.iftf.org/2006/12/prediction_mark.html.

10 Croxson and Reade (2010) study Betfair's major football markets, finding prices update remarkably quickly to the arrival of goals during live matches. The incorporation of relevant information may be slower than this in less liquid markets. Investigating horse-racing markets, Smith *et al.* (2006) have found Betfair prices to be less biased than those from bookmakers.

11 See Hanson's blog, 'Overcoming Bias' for a discussion of this initiative: www.overcomingbias.com/2006/11/first_known_bus.html.

12 The current recession may have influenced utilization of business prediction markets, but not obviously negatively: 665 of the total 1,695 respondents in the McKinsey survey felt that the economic downturn had increased interest in Web 2.0 technologies within their organizations, 704 considered it to have had no impact, and 187 reported that it had decreased interest in these tools.

13 I am grateful to McKinsey & Company for providing access to their proprietary detailed survey data beyond the summary results published at their website (www.mckinseyquarterly.com), and to Michael Chui at the McKinsey Global Institute for several helpful discussions on this topic.

14 To date, the most high-profile initiative to create a public sector information market has been a proposed information market for the US intelligence community – the 'Policy Analysis Market', which was the brainchild of academic economist Robin Hanson. With the blessing and seed funding of the US Defense Advanced Research Projects Agency (DARPA), 'PAM' was designed to aggregate information regarding geopolitical risks and terror attacks. It failed in a storm of controversy, largely for political reasons (Hanson, 2006a). The author is aware of selected more modest attempts to introduce information markets into public sector environments, including a recent experiment in the United Kingdom to deploy information markets to forecast demand for bed space at a large public hospital (Rajakovich and Vladimirov, 2009).

15 Tziralis and Tatsiopoulos (2007) provide a comprehensive survey of the prediction markets literature.

16 Lewis Shephard, Director of Microsoft's Institute for Advanced Technology in Governments, has commented on his blog that the evidence regarding prediction markets is not established enough to justify their deployment by governments:

Microsoft Research has explored prediction markets, running an internal one as the 'Information Forecasting Exchange' from 2003–2006. Internal efforts at Yahoo and Google have also been noted. But, frankly, I'm not actively promoting PM's to government friends, as I don't believe we understand the results and supporting science well enough yet.

(http://lewisshepherd.wordpress.com/2008/07/12/test-for-prediction-markets/)

17 See Levitt and List (2007) for a recent consideration of factors affecting the generalizability of laboratory findings.

18 Many other design choices must be made when implementing a prediction market, including which trading mechanism to use. See Tziralis and Tatsiopoulos (2007) for a recent overview of the sub-literature dealing with market modelling and design.

19 Wolfers and Zitzewitz (2006) have highlighted the problem of attracting non-informed traders as one of five open questions about prediction markets:

> Counterintuitively, the problem for most prediction markets is attracting sufficient uninformed order flow. Markets need uninformed order flow to function; when trading is conducted by rational traders, whose sole motivation is expected returns, the no-trade theorem binds, and the market unravels. Uninformed order flow can have a variety of motivations (entertainment, overconfidence, and hedging, for example), but with the exception of hedging, these are usually noneconomic, putting economists at a comparative disadvantage in predicting which markets will succeed.

20 'Potential traders are individuals who have information related to our private markets and include those in the healthcare and public health fields, such as physicians, nurses, microbiologists, epidemiologists, and public health professionals, among others.' http://iehm.uiowa.edu/iehm/content/faq.html.

21 Scott Page has demonstrated formally that the wisdom of the crowd depends not only on the abilities of the people within it, but also on their cognitive differences (Page, 2007).

22 In order to operate the Iowa Election Markets legally, Iowa academics obtained no-action letters from US regulator the Commodity Futures Trading Commission. To secure this relief it was agreed to limit positions to $500 and to operate on a non-for-profit basis.

23 Bo Cowgill, until recently the manager of Google's internal prediction markets, observed that Google employees seem more concerned with status than cash remuneration: 'on a number of occasions, I've forgotten to pay out the small cash prizes we have at Google, and nobody noticed. But everyone notices when the T-shirts that show who won don't come' (Dye, 2008).

24 Adam Siegel, CEO of prediction markets provider Inkling, sees this loss of control over sensitive information as a prohibitive concern for some organizations: 'Some people are simply scared of exposing sensitive information – it's too politically toxic in their organizational climate' (Comment made during interviews conducted by the author in 2007–2008.)

25 The Golem effect designates the opposite phenomena, whereby low expectations encourage low performance. Professor Dov Eden at Tel Aviv University has confirmed both self-fulfilling phenomena in banks, schools and the military, among other settings (Eden, 1984). His advice to leaders: 'Have high expectations and reinforce them with positive messages to the employee, even if it requires being a good actor' www.aftau.org/site/News2?page=NewsArticle&id= 6927.

26 Managing expectations about staff retention may be a particular concern around the time of a major organizational change, such as following a merger. According to press reports, 120 of Cadbury's 170 senior managers left the confectioner in the six months following its takeover by Kraft in February 2010. www.thegrocer.co.uk/articles.aspx?page=articles&ID=211285.

27 Comments were provided during interviews conducted by the author in the period 2007–2008.

28 Similarly, it appears that Hewlett Packard previously refined its proprietary information mechanism, BRAIN, to conceal aggregates such as the current forecast (Acheson *et al.*, 1997)

29 *Predictocracy* can be read online as a blog: http://predictocracy.org/blog/?p=107.

30 Mat Fogarty at prediction markets technology provider Crowdcast has expressed the view that, at least in the corporate setting, where play-money incentives linked to modest prizes are the norm, fears about manipulation seem exaggerated: 'Sometimes people ask about manipulation. With the current low level of prizing it is not a concern. I have not come across manipulation in practice.' These comments were provided during interviews conducted by the author in the period 2007–2008.

31 Comments were provided during interviews conducted by the author in the period 2007–2008.

32 http://blogs.hbr.org/hbr/mcafee/2009/12/prediction-markets-a-teaching-moment.html.

33 Thomas Davenport has urged organizations to shift to a 'test-and-learn mind-set', basing their decisions on the results of randomized internal experiments. See his recent article in the *Harvard Business Review* (February 2009): 'How to Design Smart Business Experiments', available online at: http://hbr.org/2009/02/how-todesign-smart-business-experiments/ar/1.

References

Abramowicz, M. (2007), *Predictocracy: Market Mechanisms for Public and Private Decision Making*, Yale University Press.

Acheson, D., R. Mazzilli and L. Prata (1997), 'Prediction Markets', mimeo, MIT Sloan Management.

Armstrong, J.S. (2006), 'How to Make Better Forecasts and Decisions: Avoid Face-to-Face Meetings', *Foresight: The International Journal of Applied Forecasting*, 5, 3–15.

Arrow, K.J., S. Sunder, R. Forsythe, R.E. Litan, M. Gorham, E.W. Zitzewitz, R.W. Hahn, R. Hanson, D. Kahneman, J.O. Ledyard, S. Levmore, P.R. Milgrom, F.D. Nelson, G.R. Neumann, M. Ottaviani, C.R. Plott, T.C. Schelling, R.J. Shiller, V. L. Smith, E.C. Snowberg, C.R. Sunstein, P.C. Tetlock, P.E. Tetlock, H.R. Varian and J. Wolfers (2007), 'Statement on Prediction Markets', *AEI-Brookings Joint Center Related Publication No. 07–11*.

Bell, T. (2009), 'Private Prediction Markets and the Law', *The Journal of Prediction Markets* 3(1), 89–110.

Berg, J., R. Forsythe, F. Nelson and T. Rietz (2008), 'Results from a Dozen Years of Election Futures Market Research', *in* C.R. Plott and V.L. Smith (eds), *Handbook of Experimental Economics*, Elsevier.

Camerer, C. and K. Weigelt (1991), 'Information Mirages in Experimental Asset Markets', *Journal of Business* 64, 463–493.

Chen, K.-Y., B.A. Huberman and L.R. Fine (2001), 'Forecasting Uncertain Events with Small Groups', *Proceedings of the 3rd ACM Conference on Electronic Commerce (EC-2001)*, Tampa, Florida, USA, 14–17 October.

Chen, Y., C. Chu, T. Mullen and D.M. Pennock (2005), 'Information Markets vs. Opinion Pools: an Empirical Comparison', *Proceedings of the 6th ACM Conference on Electronic Commerce*, Vancouver, BC, Canada, June 2005.

Chordia, T., R. Roll and A. Subrahmanyam (2008), 'Liquidity and Market Efficiency', *Journal of Financial Economics* 87, 249–268.

Christiansen, J. (2007), 'Practical Experiments in Small Markets and Behaviours Observed', *The Journal of Prediction Markets* 1, 17–41.

Clemen, R.T. and R.L. Winkler (1999), 'Combining Probability Distributions from Experts in Risk Analysis', *Risk Analysis* 19(2), 187–203.

Cowgill, B., J. Wolfers and E. Zitzewitz (2009), 'Using Prediction Markets to Track Information Flows: Evidence from Google', *Working Paper*.

Croxson, K. and J. Reade (2010), 'Information and Efficiency: Goal Arrival in Soccer Betting', mimeo, Oxford University.

D'Amuri, F. and J. Marcucci (2009), *'Google it!' Forecasting the US Unemployment Rate with a Google Job Search Index*, Institute for Social and Economic Reform.

Davenport, T.H. (2009), 'How to Design Smart Business Experiments', *Harvard Business Review*, 87, 68.

De Long, J.B., A. Shleifer, L.H. Summers and R.J. Waldmann (1990), 'Positive Feedback Investment Strategies and Destabilizing Rational Speculation', *Journal of Finance* 45, 379–395.

Dvorak, P. (2008), 'Best Buy Taps "Prediction Market"', *Wall Street Journal*, 16 September.

Dye, R. (2008), 'The Promise of Prediction Markets: A Roundtable', *McKinsey Quarterly*.

Economist, The (2005), 'Market, Market, on the Wall. Technology Trends: If Prediction Markets are so Good at Making Forecasts, Why Not Use them to Identify Emerging Technologies?', 10 December.

Eden, D. (1984), 'Self-Fulfilling Prophecy as a Management Tool: Harnessing Pygmalion', *The Academy of Management Review* 9(1), 64–73.

Forsythe, R., T. Palfrey and C. Plott (1982), 'Asset Valuation in an Experimental Market', *Econometrica* 50, 537–567.

Forsythe, R., F. Nelson, G.R. Neumann and J. Wright (1992), 'Anatomy of an Experimental Political Stock Market', *The American Economic Review* 82, 1142–1161.

Ginsberg, J., M.H. Mohebbi, R.S. Patel, L. Brammer, M.S. Smolinski and L. Brilliant (2009), 'Detecting Influenza Epidemics Using Search Engine Query Data', *Nature* 457, 1012–1014.

Graefe, A. (2009), 'Prediction Markets versus Alternative Methods: Empirical Tests of Accuracy and Acceptability', PhD Dissertation, University of Karlsruhe.

Green, K.C., J.S. Armstrong and A. Graefe (2007), 'Methods to Elicit Forecasts from Groups: Delphi and Prediction Markets Compared', *Foresight – The International Journal of Applied Forecasting*. Online: http//marketing.wharton.openn.edu/documents/research/delphi-wpv33.pdf.

Gruca, T.S., J. Berg and M. Cipriano (2008), 'Incentive and Accuracy Issues in Movie Production Markets', *The Journal of Prediction Markets* 2(1), 29–43.

Hahn, R.W. and P.C. Tetlock (2005), 'Using Information Markets to Improve Public Decision Making', *Harvard Journal of Law and Public Policy* 29, 214–289.

Hahn, R.W. and P.C. Tetlock (2007), 'Optimal Liquidity Provision for Decision Makers', *AEI-Brookings Joint Center Working Paper No. 06–18*.

Hanson, R. (2006a), 'Decision Markets for Policy Advice', *in* A.S. Gerber and E.M. Patashnik (eds), *Promoting the General Welfare: New Perspectives on Government Performance*, Brookings Institution Press.

Hanson, R. (2006b), 'Foul Play in Information Markets', *in* R.W. Hahn and P.C. Tetlock (eds), *Information Markets: A New Way of Making Decisions*, AEI Press.

Hanson, R. (2008), 'Insider Trading and Prediction Markets', *Journal of Law, Economics, and Policy* 4(2), 449–463.

Hanson, R. and R. Oprea (2007), 'A Manipulator can Aid Prediction Market Accuracy', *Economica* 76, 304–314.

Hanson, R., R. Oprea and D. Porter (2006), 'Information Aggregation and Manipulation in an Experimental Market', *Journal of Economic Behavior and Organization* 60, 449–459.

Hartzmark, S. and D. Solomon (2010), 'Efficiency and the Disposition Effect in NFL Prediction Markets', *Working Paper.*

Hayek, F.A. (1945), 'The Use of Knowledge in Society', *American Economic Review* 35(4), 519–530.

Hopman, J.W. (2007), 'Using Forecasting Markets to Manage Demand Risk', *Intel Technology Journal* 11(2), n.p.

Hussam, R.N., D. Porter and V.L. Smith (2008), 'Thar She Blows: Can Bubbles Be Rekindled with Experienced Subjects?', *American Economic Review* 98(3), 924–937.

Kumar, A. and C. Lee (2006), 'Retail Investor Sentiment and Return Comovements', *Journal of Finance* 61, 2451–2486.

LaComb, C.A., J.A. Barnett and Q. Pan (2007), 'The Imagination Market', *Information System Frontiers* 9, 245–256.

Ledyard, J., R. Hanson and T. Ishikida (2006), 'Designing Information Markets for Policy Analysis', *in* R.W. Hahn and P.C. Tetlock (eds), *Information Markets: A New Way of Making Decisions*, AEI Press.

Lei, V., C.N. Noussair and C. Plott (2001), 'Nonspeculative Bubbles in Experimental Asset Markets: Lack of Common Knowledge of Rationality vs. Actual Irrationality', *Econometrica* 69(4), 831–859.

Levitt, S. and J. List (2007), 'What do Laboratory Experiments Measuring Social Preferences Tell Us about the Real World?', *Journal of Economic Perspectives* 21(2), 153–174.

Linstone, H.A. (2007), *The Delphi Method: Techniques and Applications*, Addison-Wesley.

Lohr, S. (2008), 'Betting to Improve the Odds', *New York Times*, 9 April.

Luckner, S. (2007), 'Prediction Markets: How do Incentive Schemes Affect Prediction Accuracy?', *in* N. Jennings, G. Kersten, A. Ockenfels and C. Weinhardt (eds), *Negotiation and Market Engineering*, IBFI.

McKinsey (2009), 'How Companies are benefiting from Web 2.0: McKinsey Global Survey Results', *McKinsey Quarterly*, September.

Noeth, M., C. Camerer, C. Plott and M. Weber (1999), 'Information Aggregation in Experimental Asset Markets: Traps and Misaligned Beliefs', mimeo.

O'Brien, J. and S. Srivastava (1991), 'Dynamic Stock Markets with Multiple Assets', *Journal of Finance* 46, 1811–1838.

Ortner, G. (1998), 'Forecasting Markets – An Industrial Application, Part II', *Working Paper*, Technical University of Vienna, Department of Managerial Economics and Industrial Organization.

Ottaviani, M. and P. Sorensen (2006), 'Outcome Manipulation in Corporate Prediction Markets', *Journal of the European Economic Association* 5, 554–563.

Page, S. (2007), *The Difference: How the Power of Diversity Creates Better Groups, Firms, Schools, and Societies*, Princeton University Press.

Pennock, D.M., S. Lawrence, C. Giles and F. Nielsen (2001), 'The Real Power of Artificial Markets', *Science* 291(5506), 987–988.

Plott, C. and K.-Y. Chen (2002), 'Information Aggregation Mechanism: Concept, Design, and Implementation for a Sales Forecasting Problem', *Working Paper, No. 1131*, California Institute of Technology, Division of the Humanities and Social Sciences.

Plott, C. and S. Sunder (1982), 'Efficiency of Experimental Security Markets with Insider Information: An Application of Rational Expectations Models', *Journal of Political Economy* 90, 663–698.

Plott, C. and S. Sunder (1988), 'Rational Expectations and the Aggregation of Diverse Information in Laboratory Security Markets', *Econometrica* 56(5), 1085–1118.

Polgreen, P.M., Y. Chen, D.M. Pennock and F.D. Nelson (2007), 'Using Internet Searches for Influenza Surveillance', *Clinical Infectious Diseases* 47, 1443–1448.

Prendergast, C. (1993), 'A Theory of "Yes Men"', *American Economic Review* 83(4), 757–770.

Raban, D.R. and D. Geifman (2009), 'Designing Online Information Aggregation and Prediction Markets for MBA Courses', *Interdisciplinary Journal of E-Learning and Learning Objects: IJELLO Special Series of Chais Conference 2009 Best Papers* 5, 247–262.

Rajakovich, D. and V. Vladimirov (2009), 'Prediction Markets as a Medical Forecasting Tool', *The Journal of Prediction Markets* 3(2), 78–106.

Rhode, P.W. and K.S. Strumpf (2008), 'Manipulating Political Stock Markets: A Field Experiment and a Century of Observational Data', *Working Paper*, June.

Rosenbloom, E.S. and W. Notz (2006), 'Statistical Tests of Real-Money versus Play-Money Prediction Markets', *Electronic Mar*kets 16(1), 63–69.

Sadka, R. and A. Scherbina (2007), 'Analyst Disagreement, Mispricing, and Liquidity', *Journal of Finance* 62, 2367–2403.

Scharfstein, D.S. and J.C. Stein (1990), 'Herd Behavior and Investment', *American Economic Review* 80, 465–479.

Servan-Schreiber, E., J. Wolfers, D.M. Pennock and B. Galebach (2004), 'Prediction Markets: Does Money Matter?', *Electronic Markets* 14(3), 243–251.

Siegel, S. (1961), 'Decision Making and Learning under Varying Conditions of Reinforcement', *Annals of the New York Academy of Science*, 89, 766–783.

Smith, M.A., D. Paton and L. Vaughan-Williams (2006), 'Market Efficiency in Person-to-Person Betting', *Economica* 73(292), 673–689.

Smith, V.L., G.L. Suchanek and A.W. Williams (1988), 'Bubbles, Crashes, and Endogenous Expectations in Experimental Spot Asset Markets', *Econometrica* 56(5), 1119–1151.

Spann, M. and E. Skiera (2003), 'Internet-Based Virtual Stock Markets for Business Forecasting', *Management Science* 49(10), 1310–1326.

Sunstein, C. (2006a), *Infotopia: How Many Minds Produce Knowledge?*, Oxford University Press.

Sunstein, C.R. (2006b), 'Deliberation and Information Markets', *in* R.W. Hahn and P.C. Tetlock (eds), *Information Markets: A New Way of Making Decisions*, AEI Press.

Surowiecki, J. (2004), *The Wisdom of Crowds: Why the Many are smarter Than the Few and How Collective Wisdom Shapes Business, Economies, Societies and Nations*, Little, Brown.

Tetlock, P.C. (2008), 'Liquidity and Prediction Market Efficiency', *Working Paper*.

Tetlock, P.E. (2005), *Expert Political Judgment: How Good is it? How Can we Know?*, Princeton University Press.

Tziralis, G. and I. Tatsiopoulos (2007), 'Prediction Markets: An Extended Literature Review', *The Journal of Prediction Markets* 1, 75–91.

Vaughan Williams, L. (2005), *Information Efficiency in Financial and Betting Markets*, Cambridge University Press.

Veiga, H. and M. Vorsatz (2009), 'Price Manipulation in an Experimental Asset Market', *European Economic Review* 53, 327–342.

Veiga, H. and M. Vorsatz (2010), 'Information Aggregation in Experimental Asset Markets in the Presence of a Manipulator', *Experimental Economics*.

Wolfers, J. and E. Zitzewitz (2004), 'Prediction Markets', *Journal of Economic Perspectives* 18(2), 107–126.

Wolfers, J. and E. Zitzewitz (2006), 'Five Open Questions About Prediction Markets' *in* R.W. Hahn and P.C. Tetlock (eds), *Information Markets: A New Way of Making Decisions*, AEI Press.

6 Using prediction markets in new product development

Bernd Skiera and Martin Spann

Introduction

Between 2001 and 2010, the number of internet users worldwide more than quadrupled from less than half a billion to almost two billion (www.internetworldstats.com\stats.htm). Approximately 85 percent of internet users have bought at least one product online, and as of August 2010 the largest and fastest-growing social network platform, Facebook, connects more than 500 million active users. Roughly 250 million of them visit Facebook on any given day and spend, on average, more than 80 minutes per day on the website. This wide acceptance of the internet alters product development (Dahan and Hauser, 2002), and the persistent development of successful new products remains one of the most essential challenges for companies (Crawford and Di Benedetto, 2006). Yet, new product development still remains difficult and costly (Di Benedetto, 1999). The flop rates of newly launched products have remained high over the years, often surpassing 50 percent (Urban and Hauser, 1993). Hence, even small improvements in the new product development process can have a major effect on companies' profits and competitive advantage if this flop rate is reduced.

Therefore, new methods to improve product development are of high relevance for companies. Prediction markets, also called information markets (Hahn and Tetlock, 2006) or virtual stock markets (Spann and Skiera, 2003), are such a method. They attempt to connect a group of participants together in a virtual marketplace and enable them to trade shares of virtual stocks. In prediction markets, these stocks represent a bet on the outcome of future, uncertain events, and their value depends on the realization of these events (Forsythe *et al.*, 1992; Spann and Skiera, 2003). For example, a stock may represent the predicted number of sold units of a new product (e.g., the iPhone 4G) in the first quarter after its market introduction. After the outcome of the specific event becomes known (i.e., the actual number of units sold), each share of virtual stock receives a specified cash dividend (e.g., $1 for each 1,000 product units sold). Participants in a prediction market use their own assessments about the expected event outcome and its corresponding cash dividend to derive an expected stock value and trade accordingly. For example, a participant's expectation that 100,000 iPhones would sell during the first quarter after its market introduction corresponds to a cash dividend of $100. If the current price of the corresponding stock is $95 (or $105), the stock appears undervalued (or overvalued) to this participant, so he or she should try to earn the anticipated profit of $5 by buying (or selling). The participant's information thus affects the market price through his or her trading behavior.

Such prediction markets initially were applied in the form of political stock markets (later called the Iowa Electronic Market) to predict the outcome of the 1988 US presidential election, with participation restricted to members of the University of Iowa community (for a more detailed description, see Berg *et al.* (2008) and Spann and Skiera (2003). In the ensuing two decades, prediction markets have achieved promising results for short-term forecasting tasks, such as political events (Berg *et al.*, 2008; Forsythe *et al.*, 1992), sports competitions (Luckner and Weinhardt, 2007; Servan-Schreiber *et al.*, 2004; Spann and Skiera, 2009), business events (Elberse, 2007; Foutz and Jank, 2010; Gruca *et al.*, 2003; LaComb *et al.*, 2007; Spann and Skiera, 2003) and the identification of lead users or experts (Spann *et al.*, 2009).

The theoretical foundation for prediction markets is the market efficiency attained in a competitive market through the price mechanism, which Hayek (1945) considers the most efficient instrument for aggregating asymmetrically dispersed information possessed by various market participants. Prices in efficient markets always fully reflect the available information (Fama, 1970), so the prices of virtual stocks serve as good predictors (Spann and Skiera, 2003).

The aim of this chapter is to discuss the application of prediction markets in new product development and to empirically determine factors that influence the forecasting error of prediction markets. For that reason, Section 2 discusses the possibilities for prediction markets to support the different stages of the new product development process. In Section 3 we describe an empirical study that uses prediction markets to forecast the success of new products, compare forecasting accuracy with those of expert judgments and analyze the factors that influence forecast accuracy. Section 4 summarizes the implications of the chapter.

2 Prediction markets and their use in new product development

2.1 Product development process

Prediction markets can be of use in the four key stages of the new product development process (Dahan and Hauser, 2002), namely: (1) idea generation and screening; (2) product concept development and testing; (3) product testing; and (4) product launch. The idea generation and screening stage poses the challenges of generating a sufficiently large number of concepts that contain valuable new product ideas, as well as to efficiently screen these ideas to a manageable number of promising ones (Soukhoroukova *et al.*, 2011).

In the product concept development and testing stage, consumer preferences for different new product concepts need to be evaluated, traditionally via web-based preference-elicitation tools such as conjoint analysis (Dahan and Hauser, 2002).

Online communities can test product prototypes as part of the product-testing stage. Web-based preference-elicitation tools can be applied at this stage as well. The launch of a product can be supported by accurate pre-launch demand predictions which support production and capacity planning. Further, specific product websites (e.g., for movies) can inform consumers about the product and thereby help to reduce buyer uncertainty for fairly new products.

Prediction markets can be used as an information-gathering tool to support new product development. The different opportunities of prediction markets to provide market intelligence in the new product development process are displayed in Figure 6.1, using the four distinct stages proposed above (see also Skiera and Spann, 2004; Soukhoroukova and Spann, 2005).

2.2 Idea generation and screening

In the idea generation and screening stage, an online community can be created, which is organized around existing products that are traded on a prediction market. One example is the Hollywood Stock Exchange (www.hsx.com), which runs a prediction market on the success of new movies and contains a major virtual community dealing with movie-related topics. Thus, trading on the prediction market stimulates consumers to express and discuss new product ideas, as well as new product success factors in the online community. The systematic analysis of this community can produce new product ideas. Further, participants of this prediction market can be analyzed in order to detect lead users (Spann *et al.*, 2009).

Spann *et al.* (2009) outline, in an empirical study, that prediction markets are able to identify lead users in consumer products markets. They show that lead users perform better than the average participants in a prediction market. Their conclusion is that the use of prediction markets for screening purposes can be an efficient alternative to survey-based screening. They also suggest using prediction markets for selecting those lead users who have a better understanding of the market.

Further, idea markets can be used to generate a large number of ideas, as well as to efficiently screen these ideas (Soukhoroukova *et al.*, 2011; LaComb *et al.*,

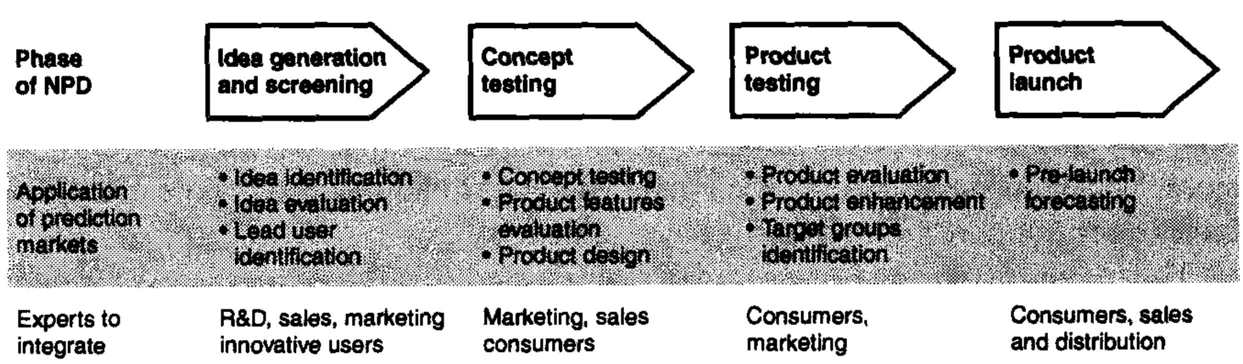


Figure 6.1 Prediction markets at key phases of new product development.

2007; Bothos *et al.*, 2009; Chen *et al.*, 2010). The main objective of idea markets is to create a virtual market in which participants can suggest new product ideas, represented as idea stocks, and collectively filter and evaluate those ideas by selling and buying idea stocks. A market mechanism adjusts the corresponding prices, which serve as indicators for the possible success of the different new product ideas. The two most important distinctions between idea markets and prediction markets are as follows. First, the initiator in traditional prediction markets determines the number of available stocks, whereas this number varies in idea markets, depending on the number of new suggestions by participants. For example, in a political stock market, the initiator would create two stocks for the 2008 US presidential election: McCain and Obama. In an idea market, however, participants can propose an unlimited and previously unknown number of different stocks that reflect their different ideas. Second, the value of the stocks in an idea market cannot depend on the realization of an actual event outcome in the near future.

Soukhoroukova *et al.* (2011) empirically explore the performance of idea markets in a real-world field study at a large, high-tech B2B (business to business) company that includes more than 500 participants from 17 countries and features various idea-sourcing tasks. Their results support the feasibility of idea markets for real-world application to support sourcing, filtering and evaluation of new product ideas.

2.3 Product concept development and testing

In the product concept stage, a preference market can try to assess consumers' aggregated preferences for different new product concepts (Dahan *et al.*, 2011, 2010). Preference markets offer a flexible prioritization methodology for product features and concepts, one that scales up in the number of testable alternatives, limited only by the number of participants. New product preferences for concepts, attributes and attribute levels are measured by trading stocks whose prices are based upon share of choice of new products and features. Dahan *et al.* (2010) developed a conceptual model of scalable preference markets, and tested it experimentally in several studies. They found that benefits of the methodology include speed (less than one hour per trading experiment), scalability (question capacity grows linearly in the number of traders), flexibility (features and concepts can be tested simultaneously) and respondent enthusiasm for the method.

One potential problem of preference markets can be the lack of an underlying value for the payoff value of stocks, because no actual sales are known at the time of concept testing. Therefore, Dahan *et al.* (2011) use the final price in the stock market as payoff value. Slamka *et al.* (2011) show that this potential limitation does not hurt the validity of the results.

The major difference between idea markets and preference markets (Dahan *et al.*, 2010) or concept markets (i.e., securities trading of concepts; Dahan *et al.*, 2011) is that the number of available stocks is governed by the initiator of the concept or preference market, whereas the participants in idea markets propose and filter the new product ideas, which are then traded. Idea markets are distinctive because they deal with an unknown number of different stocks.

2.4 Product testing

Further, different product design solutions can be evaluated in a prediction market on a company's intranet. Therefore, the assessments on the feasibility and efficiency of different construction and manufacturing solutions can be traded by one or several R&D teams. Further, the inclusion of members of the marketing department as traders can add market-related information. Prediction markets might especially be beneficial in such situations because the aggregation of the individual estimates will not be biased due to different positions in a company's hierarchy. van Bruggen *et al.* (2010) show in a different context that even prediction markets with only six participants are large enough to get good results.

Product prototypes can be tested in a prediction market so that participants can trade their assessments on the market success of these different prototypes (Dahan *et al.*, 2011). Therefore, additional information can be elicited by combining a prediction market with traditional survey and focus group methods on the same set of consumers, because trading in the prediction market can stimulate consumers to focus on the subject and quantify their assessment of market success (Spann and Skiera, 2003).

2.5 Product launch

Prediction markets can be used for pre-launch forecasting of a product's market success. Such forecasts are very useful for a company in order to optimize their product launch-related marketing instruments. For example, a movie studio can use this information to decide on promotions and advertising related to the movie's release. Movie exhibitors can plan on whether to display the movie in large or small theaters. Further, an analysis of traders' portfolios and trading behavior might be useful for the analysis of target groups.

Compared to other knowledge-gathering techniques applicable in the new product development process, prediction markets offer the following advantages (Dahan and Hauser, 2002; Spann and Skiera, 2003; van Bruggen *et al.*, 2010). First, they allow for an almost real-time reaction of stock prices to additional information and, hence, a very quick prediction of the impact of that information on future market situations. Second, it does not burden the researcher with the task of weighting and aggregating different expert judgments, as this is achieved by the trading mechanism implemented in the prediction market. Participants, for example, weight their assessments by the volume and price of the purchase or sale order they place or accept. Third, once established, a prediction market can operate at rather moderate operating costs – e.g., for repeated new product concept tests. Fourth, a prediction market provides participants with an incentive to reveal their true assessments (Forsythe *et al.*, 1999), if an adequate remuneration is properly linked to the participants' performance on the prediction market. Hence, whereas many consumer surveys remunerate consumers for their participation in a survey, a prediction market usually remunerates participants for their *successful* participation (Spann and Skiera, 2003; Dahan and Hauser, 2002). Wertenbroch and Skiera (2002) show, for example, that consumers' willingness-to-pay differs significantly according to the incentive structure being provided. Finally, participants in a prediction market might have more fun than their counterparts partaking in consumer or expert surveys (Dahan *et al.*, 2010).

3 Empirical study

The goal of the following empirical study is to analyze the use of a prediction market to predict the success of new products prior to their launch. Therefore, we analyze the feasibility, forecast accuracy and factors influencing forecast accuracy of a prediction market to predict the success of new products, namely the success of movies in Germany. Movies face high financial burdens for production and marketing, a significant failure rate and rather unstable market conditions (Sawhney and Eliashberg, 1996; Eliashberg *et al.*, 2000). Hence, we look at a prediction market that has been used as a pre-launch forecasting tool in the product launch stage.

3.1 Design of the study

We conducted the movie exchange (CMXX) seven times for the prediction of movies, using our own prediction market software (the first round also included the chart position of 11 pop-music singles in Germany, which we omit from our analysis). We conducted a prediction market for the prediction of the box-office success (number of visitors) of movies in Germany. During the seven rounds of CMXX, virtual stocks for 10–15 movies were traded in each round. In total, virtual stocks were traded for 81 movies. At the end of CMXX, each share of movie stock received a cash dividend (payoff) according to the total number of visitors of the respective movie in Germany until the end of the specified round.

Prices were limited to \$3,000 (virtual) in the first round, and \$3,500 (virtual) in the following rounds for movie stocks, considering that more than 3,000,000 and 3,500,000 movie visitors are unrealistic in Germany. In the first round, CMXX provided non-monetary incentives in the form of a "Golden Record" and ten music CDs for the participant with the highest portfolio value, and five and three music CDs for the participants with the second and third highest portfolio values, respectively. Four sets of movie merchandise were given to randomly selected participants ranking fourth to one-hundredth, according to final portfolio value. In the second to seventh rounds, the participant with the highest portfolio value in each round received an annual ticket for a large German movie exhibitor; the participants with the second and third highest portfolio value received ten free movie tickets and a set of movie merchandise, respectively. Table 6.1 provides an overview of the design of the movie exchange.

Table 6.1 Design of the movie exchange

Step	*Decisions*
Choice of forecasting goal	Forecasting the number of movie visitors in Germany Payoff function: movie visitors in Germany: €1 (virtual) per 1,000 visitors to a movie Duration: first round: 22 January–5 February 2001; second to seventh rounds: duration of one month each between May and October 2001 Open to the public; participants can join at any time
Incentives for participation and information revelation	*Composition of initial portfolios/endowment:* Endowment of 100 shares of each type of movie stock and $500,000 [$250,000] (virtual) per participant Provision of loans up to $500,000 [$250,000] (virtual) at no interest rate per participant *Remuneration/incentive mechanism:* Non-monetary rewards Rank-order tournament: rewards for participants with the highest, second highest and third highest increase in (virtual) portfolio value (annual movie ticket, ten free movie tickets, movie merchandise) Time interval: whole prediction market duration Incentives not based on performance: first round: lottery for four rewards among participants ranked fourth to one-hundredth
Financial market design	Double auction trading mechanism with open order book Trading times: 24 hours per day, seven days per week No short trading Order types: limit and market without temporal restriction No position limits, maximum price limits of 3,000 [3,500] for movie stocks No trading fee

3.2 Forecast accuracy

The price of a share of a movie stock represents a prediction of the number of visitors for the selected movie, up until the end of the specified round. Thus, by multiplying the stock price by 1,000, the forecast of a movie's number of visitors can be easily derived.

In each round, the movie exchange attracted around 50 actively trading participants. The forecasts derived from CMXX used the price of the last trade of a specific type of stock before trading was stopped at the end of a specific round. CMXX faced the problem that it included movies with very few visitors and presumably little information available among the participants (e.g., movies with as few as 20,000 visitors as well as movies with as many as 2,296,000 visitors). Consequently, forecast accuracy for the less-publicized movies with fewer than 100,000 visitors was rather poor, with an absolute percentage error of above 100 percent each (see Table 6.2).

Table 6.2 Forecast error of movie exchange

APE	*Round 1*	*Round 2*	*Round 3*	*Round 4*	*Round 5*	*Round 6*	*Round 7*	*Overall*
Movie 1	**0.024**	26.889	**0.096**	**0.050**	**0.048**	**0.012**	**0.030**	
Movie 2	2.776	**0.232**	0.061	**0.198**	**0.010**	5.000	**0.028**	
Movie 3	**0.332**	**0.031**	**0.400**	**0.130**	**0.032**	**0.252**	**0.505**	
Movie 4	0.190	**0.570**	1.667	**0.956**	**0.170**	**0.018**	**0.031**	
Movie 5	**0.057**	**0.307**	**0.123**	2.824	0.333	0.053	**0.190**	
Movie 6	**0.141**	**0.222**	**0.208**	**0.055**	**0.102**	**0.189**	n.a.	
Movie 7	9.667	**0.170**	**0.378**	**0.407**	5.600	1.000	**0.281**	
Movie 8	0.705	2.784	7.214	**0.297**	**0.153**	**0.250**	2.049	
Movie 9	3.516	**0.005**	**0.074**	**0.835**	**0.701**	n.a.	**0.040**	
Movie 10	0.630	3.839	2.258	2.390	3.115	**0.080**	**0.500**	
Movie 11	0.250	**0.263**	49.000	1.041	1.935	n.a.	**0.013**	
Movie 12		13.318		1.174		0.479	3.895	
Movie 13		**0.118**						
Movie 14		**0.133**						
Movie 15		5.507						
MAPE	*1.663*	*3.626*	*5.589*	*0.863*	*1.109*	*0.733*	*0.687*	**2.119**
Median	*0.332*	*0.263*	*0.378*	*0.621*	*0.170*	*0.220*	*0.190*	**0.263**
Min	*0.024*	*0.005*	*0.061*	*0.050*	*0.010*	*0.012*	*0.013*	**0.010**
Max	9.667	*26.889*	*49.000*	*2.824*	*5.600*	*5.000*	*3.895*	**49.000**

Notes
Bold: Movies having over 100,000 visitors; n.a.: Movie release postponed; APE: absolute percentage error.

3.3 Performance compared to expert judgments

The performance of the CMXX results is compared to corporate expert predictions from the management of a large German movie exhibitor that we were able to collect for the first two rounds, but not for additional rounds (see Table 6.3). We compared the predictions of CMXX directly. The expert predictions were provided approximately one week before the end of each round of CMXX and were not made available to the participants of CMXX. The CMXX hit rate in the first round was six out of ten for movies (for the eleventh movie the CMXX prediction and the expert prediction were identical). In the second round, the CMXX hit rate was 11 out of 15 in comparison to the expert predictions from the movie exhibitor. Table 6.3 compares the mean absolute percentage error (MAPE) of CMXX to that of the expert predictions for movies with more than 100,000 visitors. The forecasts of CMXX are significantly better than those of the experts, indicating that either CMXX performed well and/or that the experts performed poorly.

3.4 Factors influencing forecast error

The results of our empirical study demonstrate that prediction markets can sometimes produce rather weak results. Therefore, it is important to derive factors that can indicate the expected forecast accuracy of a prediction market. In this section we analyze the influence of different exogenous and endogenous factors on the forecast error of the movie exchange. Exogenous factors are the ones that are not derived from the prediction market itself, but rather depend on the product being used on the prediction market: the distribution intensity of movies in the form of the number of screens a movie is released on on opening weekend, as well as the genre of a movie (see Table 6.4). Endogenous to the stock market is the stock price volatility of a specific stock on the last five days of trading at the prediction market.

Table 6.5 displays the estimation results for the influence of endogenous and exogenous factors on the forecast error of all 81 movies traded at the movie

Table 6.3 Comparison between predictions of CMXX and experts (movies>100,000 visitors)

Instrument	*CMXX: MAPE (%)*	*Experts: MAPE (%)*	*CMXX percentage improvement (p-value)** (%)*
Round 1*	13.83	47.46	70.86 (0.331)
Round 2*	20.50	115.73	82.29 (0.010)
Round 1 and 2*	18.59	96.20	80.68 (0.005)

Notes
MAPE: Mean absolute percentage error.
* Movies having over 100,000 visitors.
** Percentage of improvement of CMXX over alternative expert judgments: =(MAPE Expert – MAPE CMXX)/MAPE Expert (two-tailed paired t-test for difference)

Table 6.4 Coding according to genre of movie

Movie genre	*Action/thriller*	*Drama/romance*	*Comedy*	*Rest*
Number of movies	22	18	26	15

Note
ANOVA (impact of genre on forecast error): F-value=1.116, p-value=0.348.

exchange. Only price volatility and the number of movie screens at the opening weekend exert a significant influence. However, these two variables (one exogenous and one endogenous) display a significant negative correlation (Pearson: –0.400 (p-value <0.01)). Thus, both the exogenous factor of the number of screens and the endogenous factor of price volatility can indicate the expected forecast accuracy of the prediction market. If we omit from all movies the 20 percent having the highest price volatility, then the mean forecast error reduces from 211.9 percent to 97.13 percent. The reduction in value for the price volatility is a coefficient of variation of 0.509 or 50.9 percent in this case. Analogously, if we omit from all movies the 20 percent having the lowest number of screens on opening weekend, then the mean forecast error reduces from 211.9 percent to 69.79 percent. The reduction in value for the number of screens is 119 in this case. The latter results are in line with the forecasting errors (71.1 percent) of Sawhney and Eliashberg (1996) in a study to predict the box-office revenues for ten movies.

4 Summary and conclusions

The results of the empirical study show that prediction markets can provide better predictions than expert judgments. Yet, there is no guarantee that prediction markets always lead to good results, and the empirical study shows that

Table 6.5 Estimation results for factors influencing forecast error of movie exchange

Parameter value (standardized)	*Model 1*	*Model 2*	*Model 3*
Constant (p-value)*	(0.008)	(0.000)	(0.613)
Price volatility (p-value)**	0.153 (0.199)		0.277 (0.012)
Number of screens (p-value)	–0.289 (0.024)	–0.319 (0.004)	
DV_Action_Thriller (p-value)	–0.252 (0.078)		
DV_Drama_Romance (p-value)	–0.227 (0.123)		
DV_Comedy (p-value)	–0.229 (0.127)		
R^2	0.171	0.102	0.077
F-value (p-value)	3.084 (0.014)	8.955 (0.004)	6.584 (0.012)

Notes
N=81 movies
* Constant: no value for standardized parameters.
** Measured as coefficient of variation.

prediction markets might also provide some rather weak forecasts. The promising result, however, is that the forecasting error might be significantly reduced by recognizing the factors that had a negative influence on forecasting accuracy in previous prediction markets. Therefore, the repeated use of prediction markets allows for developing good indicators for the expected forecast accuracy, and the price volatility might serve as a general indicator for a prediction market's predictive validity.

Prediction markets can support the new product development process at all four key stages, but may require design modifications (e.g., idea markets, preference markets). Recent studies indicate that the use of prediction markets in the new product development process is beneficial for companies. As the use of prediction markets in new product development is still an emerging research area, it appears to be a rich field for further studies.

References

Berg, J.E., Nelson, F.D. and Rietz, T.A. (2008), "Prediction Market Accuracy in the Long Run," *International Journal of Forecasting*, 24, 283–298.

Bothos, E., Apostolou, D. and Mentzas, G. (2009), "IDEM: A Prediction Market for Idea Management," in: Weinhardt, C., Luckner, S. and Stößer, J. (eds.), *Designing E-Business Systems: Markets, Services, and Networks*, Berlin: Springer, pp. 1–13.

Chen, L., Goes, P., Harris, W., Marsden, J. and Zhang, J. (2010), "Preference Markets for Innovation Ranking and Selection," *Interfaces*, 40, 144–153.

Crawford, M. and Di Benedetto, A. (2006), *New Products Management*, Boston, MA: R.D. Irwin.

Dahan, E. and Hauser, J.R. (2002), "The Virtual Customer," *Journal of Product Innovation Management*, 19, 332–353.

Dahan, E., Lo, A.W., Poggio, T., Chan, N.T. and Kim, A. (2011), "Securities Trading of Concepts (STOC)," Working Paper, University of California, Los Angeles.

Dahan, E., Soukhoroukova, A. and Spann, M. (2010), "New Product Development 2.0: Preferences Markets. How Scalable Securities Markets Identify Winning Product Concepts & Attributes," *Journal of Product Innovation Management*, 27, 937–954.

Di Benedetto, A.C. (1999), "Identifying the Key Success Factors in New Product Launch," *Journal of Product Innovation Management*, 16, 530–544.

Elberse, A. (2007), "The Power of Stars: Do Star Actors Drive the Success of Movies?" *Journal of Marketing*, 71(4), 102–120.

Eliashberg, J., Jonker, J.-J., Sawhney, M.S. and Wierenga, B. (2000), "MOVIEMOD: An Implementable Decision-support System for Prerelease Market Evaluation of Motion Pictures," *Marketing Science*, 19, 226–243.

Fama, E.F. (1970), "Efficient Capital Markets: A Review of Theory and Empirical Work," *Journal of Finance*, 25, 383–417.

Forsythe, R., Nelson, F., Neumann, G.R. and Wright, J. (1992), "Anatomy of an Experimental Political Stock Market," *American Economic Review*, 82, 1142–1161.

Forsythe, R., Rietz, T.A. and Ross, T.W. (1999), "Wishes, Expectations and Actions: A Survey on Price Formation in Election Stock Markets," *Journal of Economic Behavior & Organization*, 39, 83–110.

Foutz, N.Z. and Jank, W. (2010), "Prerelease Demand Forecasting for Motion Pictures Using Functional Shape Analysis of Virtual Stock Markets," *Marketing Science*, 29, 568–579.

Gruca, T.S., Berg, J.E. and Cipriano, M. (2003), "The Effect of Electronic Markets on Forecasts of New Product Success," *Information Systems Frontiers*, 5, 95–105.

Hahn, R. and Tetlock, P. (eds.) (2006), *Information Markets: A New Way of Making Decisions*, Washington, DC: AEI-Brookings Joint Center for Regulatory Studies.

LaComb, C.A., Barnett, J.A. and Pan, Q. (2007), "The Imagination Market," *Information Systems Frontiers*, 9, 245–256.

Luckner, S. and Weinhardt, C. (2007), "How to Pay Traders in Information Markets: Results from a Field Experiment," *Journal of Prediction Markets*, 1, 147–156.

Sawhney, M.S. and Eliashberg, J. (1996), "A Parsimonious Model for Forecasting Gross Box-Office Revenues of Motion Pictures," *Marketing Science*, 15, 113–131.

Servan-Schreiber, E., Wolfers, J., Pennock, D.M. and Galebach, B. (2004), "Prediction Markets: Does Money Matter?" *Electronic Markets*, 14, 243–251.

Skiera, B. and Spann, M. (2004), "Opportunities of Virtual Stock Markets to Support New Product Development," in: Albers, S. (ed.), *Cross-functional Innovation Management: Perspectives from Different Disciplines*, Wiesbaden: Gabler, pp. 227–242.

Slamka, C., Jank, W. and Skiera, B. (2011), "Second-Generation Prediction Markets for Information Aggregation: A Comparison of Payoff Mechanisms," *Journal of Forcasting*, forthcoming.

Soukhoroukova, A. and Spann, M. (2005), "New Product Development with Internet-based Information Markets: Theory and Empirical Application," 13th European Conference on Information Systems (ECIS), Regensburg.

Soukhoroukova, A., Spann, M. and Skiera, B. (2011), "Creating and Evaluating New Product Ideas with Idea Markets," *Journal of Product Innovation Management*, forthcoming.

Spann, M. and Skiera, B. (2003), "Internet-Based Virtual Stock Markets for Business Forecasting," *Management Science*, 49, 1310–1326.

Spann, M. and Skiera, B. (2009), "Sports Forecasting: A Comparison of the Forecast Accuracy of Prediction Markets, Betting Odds and Tipsters," *Journal of Forecasting*, 28, 55–72.

Spann, M., Ernst, H., Skiera, B. and Soll, J.H. (2009), "Identification of Lead Users for Consumer Products via Virtual Stock Markets," *Journal of Product Innovation Management*, 26, 322–335.

Urban, G.L. and Hauser, J.R. (1993), *Design and Marketing of New Products*, Englewood Cliffs, NJ: Prentice-Hall.

van Bruggen, G.H., Spann, M., Lilien, G.L. and Skiera, B. (2010), "The Performance of Small Prediction Markets and their Use in Institutional Forecasting," *Decision Support Systems*, 49, 404–416.

Wertenbroch, K. and Skiera, B. (2002), "Measuring Consumer Willingness to Pay at the Point of Purchase," *Journal of Marketing Research*, 39, 228–241.

7 Prediction market accuracy for business forecasting

Andreas Graefe

Interest in prediction markets has grown over the last decade. Using the search string "prediction markets," I searched the Google News Archive (for the number of published news articles) and Google Insights for Search (for the frequency of Internet searches) to estimate public interest in the method. The number of news articles dealing with prediction markets has constantly grown since the beginning of the decade, reaching its peak with the US presidential election in 2008. While Internet users showed no considerable interest in prediction markets before 2006 (Google provides data on Internet searches since 2004), the frequency of searches increased sharply in 2007 and, as with news articles, peaked in 2008. However, in 2009, the number of news articles and Internet searches dropped to levels of before 2006 and 2007, respectively.

The boost in the popularity of prediction markets at the beginning of the century can be traced back to two events. Ironically, the cancellation of the Policy Analysis Market (PAM) in 2003 initially introduced prediction markets to a broad public. In a review of the origin and development of the PAM project, Hanson (2007) reported that more than 600 media articles covered the story. One year later, James Surowiecki's bestselling book, *The Wisdom of Crowds*, was published, which described prediction markets as a powerful method to harness collective intelligence. In the following year, prediction markets were listed on the Gartner Hype Cycle, and the media frequently reported of companies (such as Eli Lilly, General Electric, Google, France Telecom, Hewlett-Packard, IBM, Intel, Microsoft, Siemens and Yahoo) experimenting with prediction markets.

Despite increasing interest, we do not know much about the relative accuracy of prediction markets and alternative forecasting methods. Published studies are limited and often of a small scale. In addition, prediction markets have often been compared to weak benchmarks. Here, I summarize published empirical evidence on the relative accuracy of prediction markets in the field of business forecasting.

1 Method

1.1 Literature search

Relevant studies were located through personal knowledge of the literature and examination of citations in papers of textbooks. Computer searches were not very useful. A search of the Social Science Citation Index using the search string ["accuracy" AND ("prediction markets" OR "information markets" OR "idea futures" OR "virtual stock markets" OR "decision markets")] located 13 papers, of which only the paper by Spann and Skiera (2003) was relevant for this review.

1.2 Inclusion criteria

Studies selected had to meet five criteria. First, the study was published – or accepted for publication – in a journal, book or conference proceeding. Second, the study compared the accuracy of prediction markets to a benchmark method. Third, the study analyzed forecasts from at least five independent prediction markets. Fourth, the study considered prediction market prices as forecasts by definition – that is, studies that used "models of prediction markets" were excluded. Fifth, the study analyzed problems in the field of business forecasting.

1.3 Analysis

When a study provided a number of comparisons, I used the one that represented best practice. For example, if a researcher compared 10 and 20 forecasts, I used only the comparison of the 20. When forecasts were made for different forecast horizons, I averaged across the horizons. If accuracy was assessed in terms of forecast error and hit rate, I used the forecast error. If different measures of forecast error were reported, I used the measure that was more common across studies and calculated the percentage error reduction achieved by the more accurate method to estimate relative accuracy.

2 Evaluative studies on the relative accuracy of prediction markets

As shown in Table 7.1, I found five relevant articles that included eight comparisons of prediction markets and alternative methods within the field of business forecasting. Of these, most evaluations compared prediction markets to individual and combined judgment (three comparisons each). One comparison was found for prediction markets and a naïve model, as well as an econometric model.

Table 7.1 illustrates a lack of empirical evaluations. In addition, the validity of the available evidence is limited, as many studies used similar data or analyzed similar problems. Five of the eight comparisons were conducted within the

Table 7.1 Overview of comparisons

Prediction market vs.	*Study*	*Task*	*Prediction market software*	*Criterion*	*No. markets*	*Error reduction (in %)*
Naïve model	Spann and Skiera (2003)	Usage of business services	Own software	MAPE	5	28
Econometric model	Goel *et al.* (2010)	Success of movies	HSX	RMSE	97	6
Individual judgment	Spann and Skiera (2003); Pennock *et al.* (2001)	Success of movies	HSX	MAPE/MAE	140/24/109	2
Individual judgment	Pennock *et al.* (2001)	Oscar Award winners	HSX	Scoring rule	8	1
Individual judgment	Van Bruggen *et al.* (2010)	Financial indices	Own software	MAPE	20	–12
Staticized group	Gruca *et al.* (2003)	Success of movies	IEM software	MAPE	8	9
Staticized group	Pennock *et al.* (2001)	Oscar Award winners	HSX	Scoring rule	8	–2
Staticized group	Van Bruggen *et al.* (2010)	Financial indices	Own software	MAPE	20	–29
						Mean = 0

movie industry (three for predicting the box-office success of movies and two for predicting Oscar Award winners). Four of these five comparisons used data from the Hollywood Stock Exchange (HSX). Note that I combined the findings from two studies (Pennock *et al.*, 2001; Spann and Skiera, 2003), as their samples might have overlapped. Of the eight comparisons, only three were conducted outside of the movie industry, and only one was conducted within an actual organization (a small-scale comparison of prediction markets and naïve models).

The lack of empirical evidence on the relative accuracy of prediction markets for business forecasting is a major conclusion from this review.

3 Evidence on the relative accuracy of prediction markets

3.1 Vs. naïve models

Naïve models assume that things remain the same as they have been in the past. Thus, they are commonly used as benchmark models in situations involving high uncertainty. For example, for time-series forecasts, the naïve model forecast would predict the latest observation as the new forecast (no-change model).

Spann and Skiera (2003) reported results from a small-scale study, in which 20 employees of the marketing and planning departments of a large German mobile phone operator were invited to participate in a play-money prediction market. The goal was to forecast the usage of five specific mobile phone services. Twelve participants made at least one transaction.

The authors compared the prediction market results to forecasts with four simple extrapolation models (arithmetic mean, geometric mean, linear trend and exponential trend). Extrapolation models (also referred to as univariate time-series forecasting) assume that the variable being forecast will continue to behave as it did in the past. Note that extrapolation models are more sophisticated than the simple naïve model. They rely on quantitative procedures to analyze historical values of the variable being forecast. Extrapolation models are useful if many forecasts are needed; if no substantial changes are expected in the trend; if the historical trend is long; and if the historical data are reliable and valid. For more information on the conditions for using extrapolation models, see Armstrong (2001a).

I classified the extrapolations used by Spann and Skiera (2003) as naïve forecasts since they were based on very few data points (i.e., three or five observations per forecast). In addition, the comparison favored the prediction market as the study used the last traded market price, whereas the last data point for the extrapolation models was obtained a few weeks earlier.

Across the five observations, the prediction market forecasts were more accurate than each of the four extrapolation models, with an MER (mean error reduction) ranging from 12 percent to 38 percent. Compared to the typical model, the prediction market reduced forecast error by 28 percent.

3.2 Vs. econometric models

Econometric models rely on statistical (regression) techniques to estimate model parameters from historical data. Models consist of one or more regression equations used to represent the relationship between a dependent variable and explanatory variables. Econometric models are useful in situations with few variables and many reliable observations, where the causal factors vary independently of one another. Important principles for developing econometric models are to: (1) use prior knowledge and theory, not statistical fit, for selecting variables and for specifying the directions of effects; (2) use simple models; and (3) discard variables if the estimated relationship conflicts with theory or prior evidence (Green *et al.*, 2010).

Goel *et al.* (2010) compared the relative accuracy of forecasts from the HSX and an econometric model for predicting the opening week box-office revenues for 97 movies. This model drew upon prior research, which showed that the number of screens on which a movie opens and the number of Internet searches for the movie in the week before its opening are useful predictors of box-office revenue. The model forecasts were calculated by $N-1$ cross-validation. That is, the authors used the observations from 96 movies to build the model for predicting the box office revenue of the one remaining observation.

Overall, the prediction market yielded a forecast error that was 6 percent lower than the error of the econometric model.

3.3 Vs. individual judgment

If available data are inadequate for quantitative analysis or if qualitative information is likely to increase the accuracy, relevance or acceptability of forecasts, one way to make forecasts is to ask experts to think about a situation and predict what will happen. This approach is fast, inexpensive when few forecasts are needed, and may be appropriate when small changes are expected. It is most likely to be useful when the forecaster knows the situation well and gets good feedback about the accuracy of his forecasts (e.g., weather forecasting, betting on sports and bidding in bridge games).

Spann and Skiera (2003) analyzed the relative accuracy of the HSX and individual judgment. The task was to predict box-office receipts at the opening weekend of new movies that were released between January 2000 and May 2001. The authors compared the accuracy of the HSX forecasts to individual predictions from two experts. For a sample of 24 movies, the HSX yielded a MAPE that was 24 percent lower than the MAPE derived from weekly predictions published at a movie website (i.e., the first expert). Comparisons to forecasts from a second expert, the movie columnist Brandon Gray of Box Office Mojo (BOM), were possible for 140 movies. BOM yielded a forecast error that was 11 percent lower than the forecast error of the HSX. Pennock *et al.* (2001) reported results from a similar study, in which the authors analyzed the performance of the HSX for predicting the opening box-office receipts for 109 movies that were released between March 2000 and September 2000. As Spann and Skiera (2003), the authors also used the BOM expert forecasts as a benchmark for the accuracy of the HSX. Again, BOM was more accurate, yielding a forecast error that was 7 percent lower than the error of the HSX. However, note that Pennock *et al.* (2001) and Spann and Skiera (2003) analyzed movies from the same time period and, thus, their samples might overlap. For my analysis I combined the error reduction from the three comparisons. On average, the prediction market forecasts yielded a forecast error that was 2 percent lower than the forecast error achieved from individual judgment.

Pennock *et al.* (2001) reported results from a second small-scale study, in which they compared the accuracy of the HSX to individual forecasts from five movie columnists for predicting Oscar Award winners in 2000. The forecasts of experts and the HSX were obtained on the same day. The data revealed that the HSX was 1 percent more accurate than the typical individual expert. The error of the typical expert is the error one would get if one would randomly pick the forecast of one of the five movie columnists.

Van Bruggen *et al.* (2010) conducted experiments to compare the relative accuracy of prediction markets and individual judgment for forecasting two financial indices (i.e., the Dow Jones Index and the Crude Oil Spot Market Price). Participants were 60 business school students who were assigned to groups of six to participate in ten small prediction markets. From each group, the authors collected the individual judgment from the one group member with the highest knowledge score (based on responses from prior knowledge questions). Over all 20 forecasts, on average, the prediction markets were 12 percent less accurate than individual judgment.

3.4 Vs. staticized groups

Under most circumstances, the combination of several individual forecasts will be more accurate than the typical individual forecast (Armstrong, 2001b). Two studies compared the accuracy of prediction markets to statistically combined forecasts (mean or median) from a group of individuals.

Gruca *et al.* (2003) used the IEM software for predicting the box-office performances of eight movies during the first four weeks the movies were shown in theaters. Participants were mostly MBA students; no information was given about the number of participants. Trading began from 4–14 days before the opening of the movies in the theaters. The last traded market price at midnight before the opening day of each movie was used as the prediction market forecast. These forecasts were compared to median group estimates, which were derived from a survey of the prediction market participants. The setting favored prediction markets as the survey was conducted before the market opened for trading and, thus, could not incorporate information that became available during the last days before the movies were first shown. Across the eight movies, the prediction markets yielded a forecast error that was 9 percent lower than the median group estimate. However, the median group estimate was obtained up to two weeks earlier.

In their study of forecast accuracy for predicting Oscar Award winners, Pennock *et al.* (2001) also reported the average group forecast of the five movie columnists. On average, the error of the combined forecast was 2 percent lower than the error derived from the prediction markets.

Based on their experimental results, Van Bruggen *et al.* (2010) compared the relative accuracy of the prediction market results and the combined (mean) judgment of all group members. On average, the staticized groups were more accurate, yielding a forecast error that was 29 percent lower then the error derived from the prediction market forecasts.

4 Discussion

The limited empirical evidence available to date draws a mixed picture of the relative accuracy of prediction markets compared to alternative forecasting methods in the field of business forecasting. While this review found prediction markets to be more accurate than naïve and simple econometric models, gains in accuracy compared to individual judgment were small. Two comparisons found prediction markets to be (slightly) more accurate than individual judgment, whereas one comparison found individual judgment to be more accurate.

In general, one would expect prediction market forecasts to be advantageous compared to individual judgment. The reason for this goes back to the well-established principle of combining (Armstrong, 2001b), whereupon the combination of several individual forecasts will be more accurate than the typical individual forecast under most circumstances. Simply put, prediction markets provide another means to aggregate and combine information from groups.

However, this review found little advantage of prediction markets compared to the most straightforward way of soliciting information from groups: ask individuals to provide forecasts without interacting, and statistically combine the individual estimates. Of the three studies that compared prediction market forecasts and staticized groups, two found prediction markets to be less accurate, whereas one found prediction markets to be superior.

Since prediction markets seem to offer little advantage over simple group averages, the method stands as one of a number of methods for improving accuracy. Thus, the choice among which method to use would depend upon the costs and acceptability of the various methods. In order to solicit and combine individual estimates, a decision-maker needs little more than access to email and a calculator. By comparison, it is more difficult and also more expensive to launch a prediction market. One either needs to develop and design prediction market software or purchase the service from a prediction market vendor.

An additional barrier for the implementation of prediction markets in practice might be disaffection with the method. In their laboratory experiments, Graefe and Armstrong (2011) asked participants about their perceptions of each of four group interaction methods (prediction markets, the Delphi method, nominal groups and traditional face-to-face meetings). Participants rated methods involving personal communication (i.e., meetings and nominal groups) more favorable than the computer-mediated Delphi and prediction markets. In particular, participants in meetings and nominal groups experienced higher levels of cooperation in their groups and perceived group interaction as more effective. Prediction markets were rated least favorable. Prediction market participants were least satisfied with the group process and rated the method highest in terms of difficulty of participation. This is not surprising as it is not intuitive to reveal one's information by the process of trading.

To advance the field and to transfer prediction markets to practical use within organizations, we need an increasing number of empirical studies that evaluate the relative accuracy of prediction markets for business forecasting. Future research should focus on specific conditions under which organizations could benefit from prediction markets. For example, prediction markets should be valuable in situations where dispersed information becomes frequently available, as the market could continuously incorporate such information. In contrast, asking experts to reveal individual forecasts, participate in a Delphi or attend a meeting are one-off activities that need to be triggered. Prediction markets might have aggregated information by the time a decision-maker recognizes the necessity to obtain information. For a similar reason, prediction markets might be useful if one needs many forecasts that need to be continuously updated. In these cases, the initial costs for setting up the market might pay off in the long run. Also, prediction markets should be valuable in solving complex problems where information is widely dispersed among people. The results from the experiment by Van Bruggen *et al.* (2010) suggest that prediction markets have advantages over simple group averages in situations where information is unequally distributed among people.

5 Conclusion

This review summarized published empirical evidence on the relative accuracy of prediction markets and alternative forecasting methods for the field of business forecasting. Over all eight comparisons, there were no differences in the relative accuracy of prediction markets and alternative methods. While prediction markets were more accurate than naïve and simple econometric models, evidence on the relative performance of prediction markets and individual and combined judgment was mixed.

The lack of evaluation studies is one of the major conclusions of this review. Future research should further evaluate the relative accuracy of prediction markets to identify conditions under which the method is favorable to alternative forecasting methods. The question of whether prediction markets are more accurate than alternative approaches is an empirical one.

References

Armstrong, J.S. (2001a). Extrapolation for time-series and cross-sectional data. In J.S. Armstrong (Ed.), *Principles of forecasting. A handbook for researchers and practitioners* (pp. 217–243). Norwell: Kluwer.

Armstrong, J.S. (2001b). Combining forecasts. In J.S. Armstrong (Ed.), *Principles of forecasting. A handbook for researchers and practitioners* (pp. 417–439). Norwell: Kluwer.

Goel, S., Reeves, D.M., Watts, D.J. and Pennock, D.M. (2010). Prediction without markets. *EC'10 Proceedings of the 11th ACM conference on Electronic Commerce*, June 2010, ACM New York, NY, USA.

Graefe, A. and Armstrong, J.S. (2011). Comparing face-to-face meetings, nominal groups, Delphi and prediction markets on an estimation task. *International Journal of Forecasting*, 27, 183–195.

Green, K.C., Graefe, A. and Armstrong, J.S. (2010). Forecasting principles. In M. Lovric (Ed.), *International Encyclopedia of Statistical Science* (pp. 527–534). Heidelberg: Springer.

Gruca, T.S., Berg, J.E. and Cipriano, M. (2003). The effects of electronic markets on forecasts of new product success. *Information Systems Frontiers*, 5, 95–105.

Hanson, R. (2007). The policy analysis market: a thwarted experiment in the use of prediction markets for public policy. *Innovations: Technology, Governance, Globalization*, 2, 73–88.

Pennock, D.M., Lawrence, S., Nielsen, F.A. and Giles, C.L. (2001). Extracting collective probabilistic forecasts from web games. *Proceedings of the Seventh ACM SIGKDD International Conference on Knowledge Discovery and Data Mining* (pp. 174–183). New York.

Spann, M. and Skiera, B. (2003). Internet-based virtual stock markets for business forecasting. *Management Science*, 49, 1310–1326.

Van Bruggen, G.H., Spann, M., Lilien, G.L. and Skiera, B. (2010). Prediction markets as institutional forecasting support systems. *Decision Support Systems*.

8 Price biases and contract design

Lessons from Tradesports

Richard Borghesi

1 Introduction

To understand why prediction markets are valuable, one must first recognize that the market price of any financial asset contains a great deal of information. By simply observing stock price, for example, one develops a clear understanding of the future profitability and risk of a firm. For instance, if a stock's *price* is \$100, then we expect that the present *value* of all of that stock's future cash flows is also \$100. To independently arrive at the \$100 value, one would have to synthesize a great deal of information, including the estimated size and timing of all future dividend payments, plus capital gains, and the estimated risk associated with each cash flow. Instead, by simply observing market price, one can infer a great deal.

A mountain of financial research (e.g., Gruber 1996; Rubinstein 2001; Schwert 2001) shows that the vast majority of all stocks trade at prices that are extremely close to their values. This is because there is generally a sufficient number of traders in the market who are sophisticated enough to accurately determine each stock's true value, and then buy (if the price is less than the value) or sell (if the price is greater than the value). Such buying and selling activity ultimately forces prices very close to values. When prices equal values, the market is said to be price efficient, or in a state of equilibrium.

Price is generally readily observable, especially in the case of publicly traded stocks. However, stock value is often difficult, or more likely impossible, to determine. For example, the price of Enron's stock peaked at \$90.56 in August 2000, just over a year before the firm's collapse. In retrospect, its value was far less than its price, though it would have been quite difficult to determine this based on publicly available information at the time. The fact that the true value of stocks is never revealed is the genesis of the *joint hypothesis problem*. That is, in testing whether a particular stock or trading strategy has earned an abnormal rate of return, one is actually simultaneously testing whether returns are abnormal *and* whether the model correctly identifies true value.

For years, economists have examined sports-betting markets as a laboratory to better understand trader behavior and price efficiency in stock markets while avoiding the joint hypothesis problem. The two markets are close analogs because each is characterized by large amounts of public and private information, many participants, competitive bidding, transactions costs and the presence of market professionals. The advantage of examining sports-betting markets is that values of bets are revealed once the underlying contest is over. At that point, one can simply observe the outcome (by how many points one team defeated another) and measure, for instance, how frequently favorites beat underdogs, on average.

However, there are important differences between stock and sports-betting markets that make comparisons problematic. In stock markets, information is flowing constantly, and prices change rapidly in response. In traditional sports-betting markets (e.g., casinos), bets must be placed before the sporting event begins, and it is only after the event begins that information regarding the true value of the bet begins to arrive. So, in traditional betting markets, there is no way to test the ability of traders to quickly and accurately update their prior estimates of value in response to news innovations.

Tradesports, however, allowed participants to bet on the outcomes of sporting events while those events were active. For example, one could bet on an American football game after kickoff and up to the final whistle. Before kickoff, information is flowing very slowly, so comparisons to actively traded stocks are less meaningful. After kickoff, however, the constantly changing field position, score and injury status of key players strain the participants' abilities to quickly and accurately re-value the traded assets and buy or sell accordingly. Thus, this and similar markets are useful tools to researchers. Unfortunately, legal problems have forced many online exchanges to close operations within the United States.

In the United States the legality of online prediction markets such as Tradesports.com has been debated, and while the issue remains somewhat unresolved, at least one prosecution has resulted. In March 2007, Gary Kaplan, the founder of BetOnSports.com, was arrested and charged with several offenses. Chief among these was violating the Federal Wire Act of 1961. This law prohibits the use of interstate telecommunications to place wagers. In 2009 Kaplan received a four-year prison sentence.

To date, the US government has targeted just those who *operate* betting websites; no gamblers have been prosecuted. Some organizations, such as Intrade.com, that operate outside the United States have continued to accept bets from those inside US borders. In addition, the US government has allowed the Iowa Electronic Markets to operate freely.

In November 2008, Tradesports ceased its sports operations, but continued to offer non-sports-related contracts via Intrade.com. Tradesports' demise was largely driven by passage of the US Congress' Unlawful Internet Gambling Enforcement Act (UIGEA), which prohibits the transfer of funds from US financial institutions to almost all internet gambling sites. Enactment of this legislation increased the burden borne by Tradesports participants in funding their accounts, and resulted in a significant decline in trading, which ultimately forced exchange closure.

Tradesports, despite its ultimate demise, remains perhaps the best laboratory for learning about the information-processing abilities and behavior of traders. Furthermore, the massive amount of data available from Tradesports' archived sports contracts, along with the diversity of contract formats and characteristics, teaches us much about the optimal design of contracts traded on prediction markets in general.

2 What we have learned from Tradesports

Tradesports contracts represented binary options with payouts of either $0 (a "loss") or $100 (a "win") for each ten-contract lot. Real (as opposed to play) money was used. Several different classes of contracts were offered, including financial-, political- and sports-related. The outcomes of its sports contracts were relatively unambiguous. However, other categories of contracts were the object of dispute. For example, Tradesports listed a contract that was intended to expire at $100 if North Korea successfully test-fired a ballistic missile that landed outside its airspace, and at $0 otherwise. Tradesports specified that the US Department of Defense (DOD) was to be the authority relied upon to determine whether a successful launch had occurred. While many media outlets reported the success of the launch, and a variety of governments also recognized this to be the case, the DOD never released an official statement. As a result, this contract expired at zero. Such a controversy highlights the importance of precisely defining contract outcomes.

Aside from ensuring that outcomes are unambiguous, several other factors are important to the efficient functioning of prediction markets. Among these are: contract price; relative proportion of buyers to sellers; contract liquidity; the reverse favorite–longshot bias; rate of information flow; magnitude of news innovation; direction of news innovation; disposition effect; and the availability of arbitrage opportunities.

2.1 Contract price

For prediction markets having asset prices between $0 and $100, prices should translate directly into probabilities. That is, if the price is $25, then the probability of the contract expiring at $100 should be 25 percent.[1] But, surprisingly, price itself determines a considerable degree of mispricing. Wolfers and Zitzewitz (2006) show that in the Iowa Electronic Markets, prices significantly deviate from values within certain price bands. Specifically, contract win rates are lower (higher) than expected, around $25 ($75). In other words, sellers (buyers) are net winners for trades occurring at prices around $25 ($75). This pattern can be explained by the utility-maximizing behavior of traders. For prediction markets having a $0–$100 asset pricing format, and for constant relative risk aversion (CRRA) between 0 and 1, an equilibrium occurs in which price is between the market's mean subjective valuation and $50.[2] The result is that contracts costing less (more) than $50 may be overpriced (underpriced).

Figure 8.1 graphically illustrates the supply and demand for two traders whose beliefs about the probability of expiry at $100 differ. Suppose the true probability of expiry is 40 percent. One might expect that the contract price would be $40. But suppose one agent is optimistic (believes the chance of expiry at $100 is 50 percent) and the other is pessimistic (believes the chance of expiry at $100 is 30 percent). The resulting demand is 2.08 contracts, and the resulting supply is 1.92 contracts.[3] Due to this excess demand, equilibrium contract price will be above $40.

However, in the Tradesports National Football League (NFL) market, evidence shows that contracts priced in both the low- *and* mid-price range are overvalued, while those toward the upper end of the price range are undervalued. In other words, low-priced and mid-priced (high-priced) contracts win at a rate lower (greater) than expected. According to Figure 8.2, which illustrates this trend, if one were to buy a contract for a price of $50, one would have only a 42 percent chance of winning (defined as expiry at $100). That is, participants pay too much for $50 contracts. While utility-maximizing behavior explains at least

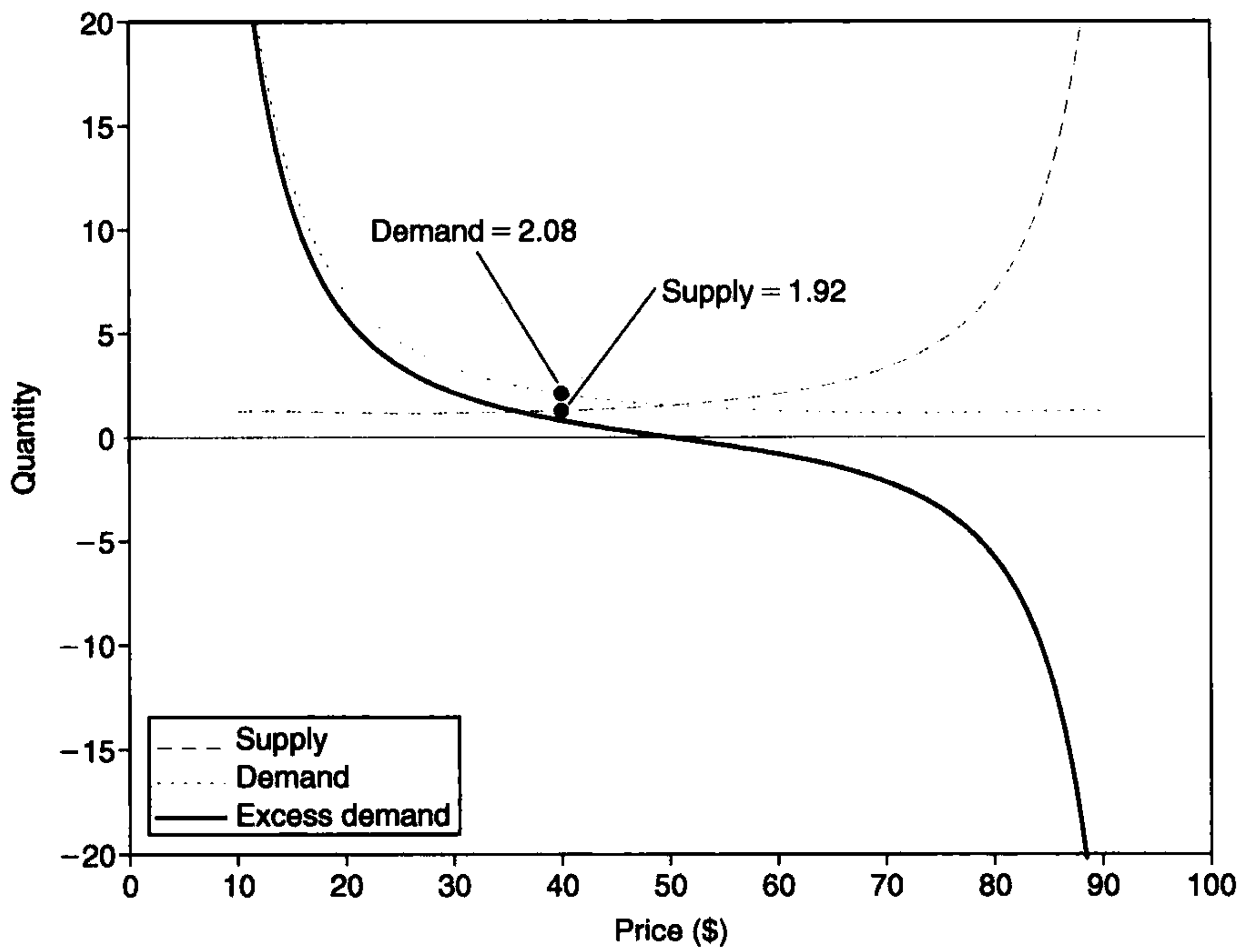


Figure 8.1 A plot illustrating the predicted differences between supply and demand under a CRRA utility model with $u(x)=x^{0.5}$. The probability of an event occurring is $p=Price/\$100$; pessimistic agent 1 has belief $b_1=p-0.10$; optimistic agent 2 has belief $b_2=p+0.10$; and each agent i has wealth $w_i=\$100$. Quantity is expressed as the number of ten-contract lots supplied and demanded at each price level.

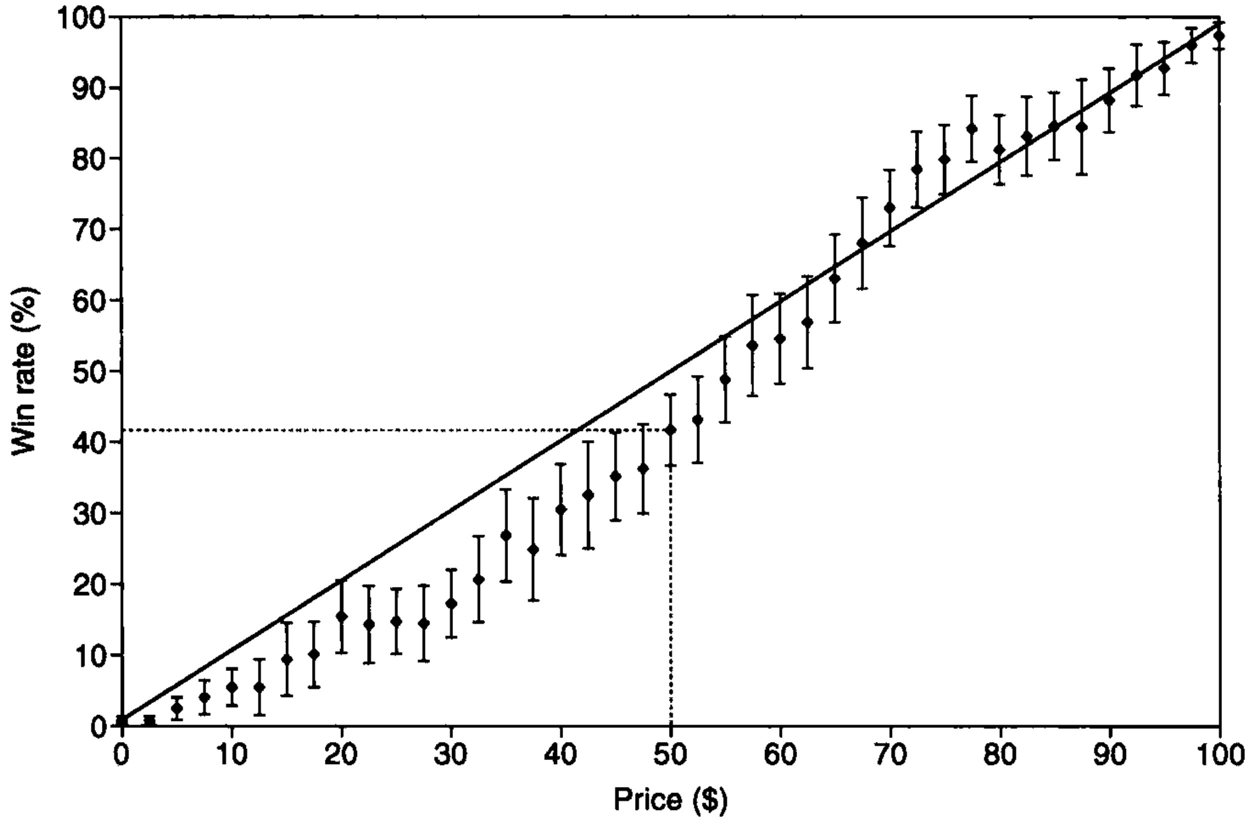


Figure 8.2 A plot generated from 216,564 NFL contract trades occurring after kickoff. Each data point represents the observed mean win rate of contracts purchased at each price level (grouped into $2.50 bands). A win is defined as expiry at $100. Error bars indicate 95 percent confidence limits.

part of the deviation between probabilities and prices at the low- and high-end of the spectrum, the mispricing of mid-price contracts is more difficult to understand, and two important questions arise from this latter observation. First, why are traders willing to pay a premium on average to buy contracts? Second, why do arbitrage operations fail to correct such a significant price anomaly? These questions are addressed in detail below.

2.2 Relative proportion of buyers to sellers

One possible explanation for the observed aggregate price premium is that the participants in this market are accustomed to buying (as opposed to selling) bets. The logic is as follows. In traditional sports-betting markets, in a match-up between Team A (the favorite) and Team B (the underdog), participants either buy a bet on Team A or else buy a bet on Team B. However, on Tradesports, all contracts are framed with respect to the favorite. Therefore, one would either buy a bet on Team A or else *sell* a bet on Team A, as there is no opportunity to buy a bet on Team B.

Suppose the contract under consideration is titled "Rams@Patriots.Patriots-11.5." The team before (after) the @ is the away (home) team, and the favored team and associated point spread are listed after the initial period. In this example, if a bettor believes that the Patriots would win by more than 11, she would buy the contract. If she believes that the Patriots would lose or win by 11 or fewer, she would have to *sell* the contract. Selling is a concept that is foreign to most, and the difficulty in calculating possible profits and losses from selling a contract likely causes many bettors to avoid taking short positions. The resulting decrease in sell-side supply creates a pressure that forces price upwards to a point above value. This is referred to as the *supply suppression hypothesis*. Buyers must spend more than fair value in order to buy the bet, and seem to be willing to do so to a certain extent.

Why do arbitrage operations not correct this inefficiency? The failure of arbitrage operations potentially can be explained by the risk associated with trading an asset that experiences frequent and significant news innovations. When the sporting event is active, massive price changes frequently occur (when big plays happen). Future play outcome is impossible to predict, so even informed bettors would be exposed to a great deal of risk in attempting to buy or sell until prices equal values. And, evidently, without the possibility of riskless profits, arbitrage operations are largely absent.

2.3 Liquidity

Market depth in the Tradesports NFL market varied greatly across assets. Some contracts had a great deal of depth (e.g., Monday night and playoff games). Others (e.g., division and conference championship contracts) had very limited depth, especially early in the season. O'Connor and Zhou (2008) study the determinants of liquidity and find that contracts on teams with better records and on those from cities having larger populations were more liquid.

However, Borghesi (2007) examines the effect of liquidity on price efficiency, and finds that for those events having the highest trading volume, overpricing is most prevalent. This potentially means that the noisy trades of unsophisticated investors overwhelm the ability of sophisticated traders to correct prices. Hartzmark and Solomon (2009) also find that mispricing is not reduced in higher-liquidity games. On the other hand, Tetlock (2004) finds no difference in returns between featured sports events (those receiving the most television network coverage) and non-featured sports events, nor between high- and low-volume contracts. But the data in the latter study are gathered by taking price snapshots at 30-minute intervals, so nearly all observations come from the inactive pre-kickoff period of trading.

So, while increased liquidity frequently means more efficiency in traditional financial markets, in the Tradesports market, this is not generally the case.

2.4 Reverse favorite–longshot bias

Prior sports-betting research has shown that the public has a strong preference for favorites in point spread sports such as football and basketball. This phenomenon is referred to as the reverse favorite–longshot bias, which proposes that bettors overestimate the likelihood that the favorite will win the game by a margin at least as great as the point spread.[4] There is no reason to believe that such a bias would be absent from the Tradesports market for NFL bets. This suggests that inferences generated from examining Tradesports sides contracts may not necessarily be applicable to non-sports-related prediction markets.

Exacerbating the potential reverse favorite–longshot bias is the fact that all Tradesports sides contracts are framed exclusively with respect to favorites. That is, if the Patriots were a stronger team than the Rams, you would not find a contract titled "Rams@Patriots.Rams+11.5." Such a contract would always be framed as "Rams@Patriots.Patriots-11.5." So, buyers must take the Patriots' side. If it is the reverse favorite–longshot bias that causes aggregate overpricing, it may be that non-sports related (e.g., internal corporate) prediction markets assets do not suffer from the price bias.

In Borghesi (2010), totals contracts are used to test whether Tradesports point spread contracts suffer from larger price biases than do its totals contracts. The expiry value of totals contracts are determined by the combined points of two teams at the end of a sports contest. An example of a Tradesports NFL totals contract is "NFL.Rams@Patriots.Over53.5." In this case, if the Rams and the

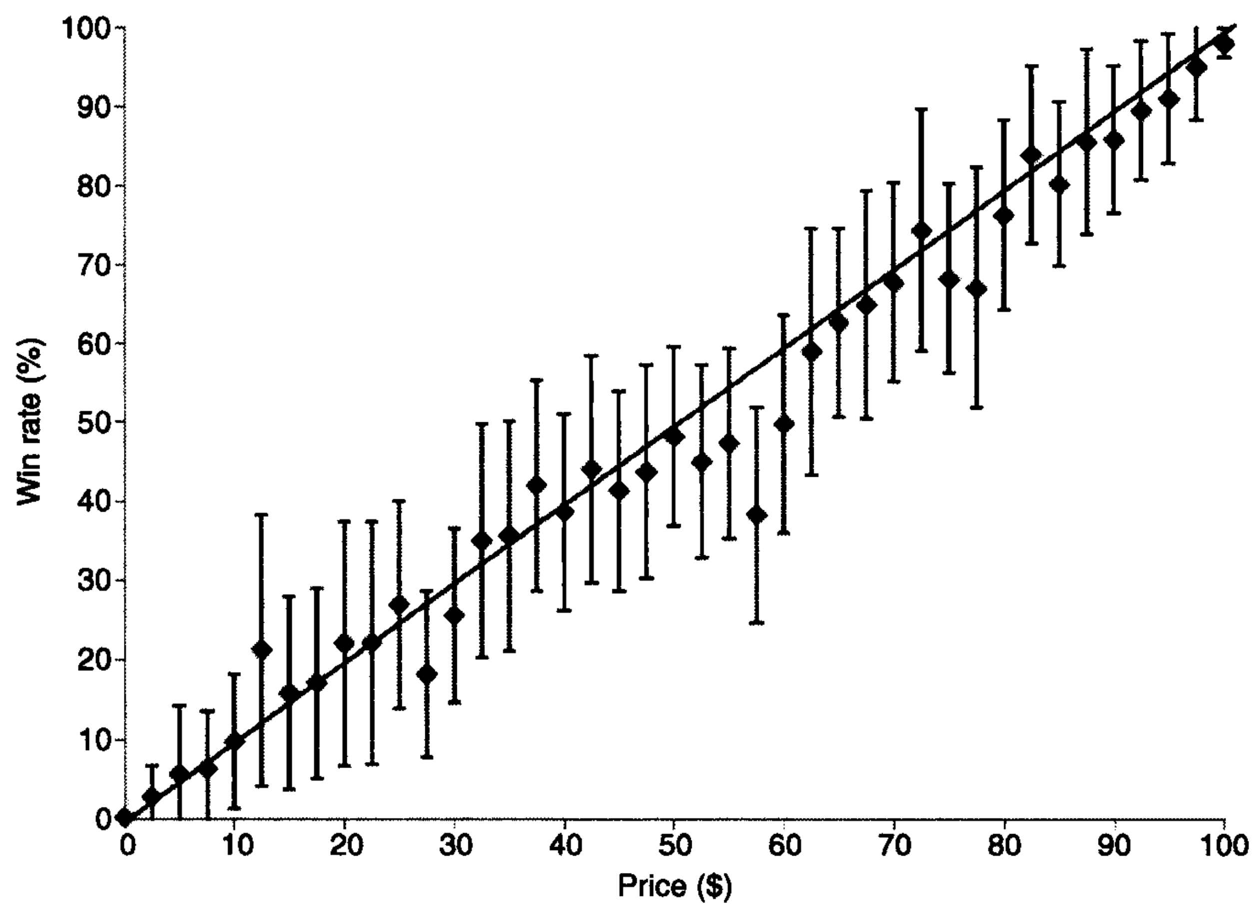


Figure 8.3 A plot generated from 33,528 NFL totals contract trades occurring after kickoff. Each data point represents the observed mean win rate of contracts purchased at each price level (grouped into $2.50 bands). A win is defined as expiry at $100. Error bars indicate 95 percent confidence limits after adjusting standard errors for the effects of clustering.

Patriots combine to score more than 53 points, the contract expires at $100, otherwise it expires at $0. Thus, there is no favorite. This study finds that while some of the biases that plague sides contracts do persist in the totals market, their magnitudes are significantly less (Figure 8.3). Specifically, the characteristic overpricing (underpricing) below $65 is muted, as is the overpricing following news innovations.

The reverse favorite–longshot bias may also extend to non-sports-related contracts where one choice is perceived to be superior to another. For example, suppose a contract is titled "Sales of Printer A will exceed Sales of Printer B." If Printer A is believed by traders to be clearly superior to Printer B, then it may be the case that bettors will overestimate the likelihood that Printer A sales will be greater. Perhaps if the true likelihood that Printer A will outsell B is 80 percent, we might expect the price of the contract to be $85, based on the known preference to take the stronger side. Potentially, this price bias could be mitigated by framing internal corporate prediction market contracts similar to totals contracts.

For example, it may be preferable to avoid listing a contract having the structure "Sales of Printer A in May will exceed sales of Printer B in May" (the analog is a sides contract in which one side is a favorite), and instead list a contract taking the form "Sales of Printer A and Printer B will total 2,000 units in May" (the analog is a totals contract).

2.5 Rate, magnitude and direction of news innovations

Two studies that focus on NFL (Borghesi, 2007) and National Basketball Association (NBA) (Borghesi, 2009) contracts teach us several important lessons. The former examines the relationship between prices and values for contracts traded both before and during active NFL games. Contrasting these two time periods sheds light on the importance of the rate of information flow on contract price efficiency. Before kickoff, information is flowing at a relatively slow rate. The few news innovations that may occur are relatively insignificant predictors of game outcome. For example, injury reports, weather reports and player disciplinary actions are some of the possible news events that occur during this time period. After kickoff, as the events of the game are unfolding, bettors are making drastic changes to their prior estimates of the likelihood of each team covering.

Evidence suggests that pre-kickoff contract prices closely resemble true contract values. Post-kickoff prices, however, deviate substantially from values, and it has been demonstrated that assets are more likely to be overpriced immediately after information shocks arrive. The conclusion to be drawn here is that during periods of rapid information flow, prediction markets may be relatively less accurate event predictors. Not only are news innovations rapidly occurring in active NFL games, but the magnitude of the arriving information is also likely to be quite large. For example, long plays, turnovers, scores, etc. are all significant determinants of game and contract outcomes. The increased outcome uncertainty during periods of rapid information flow exacerbates the hesitancy of potential sellers, and further contributes to the relative imbalance between the number of buyers and the number of sellers.

This stands in contrast to relatively mild events that occur within NBA games; the most significant events in basketball games are relatively less impactful on contract outcome. That is, the number of scoring plays in the NBA far surpasses that in NFL games, but the impact of each upon outcome is far less. No single plays (with the exception of those occurring in the final few seconds) have an impact as significant as that of turnovers or unexpected touchdowns in a football game. So, any observed differences between the price efficiency of NFL and NBA contracts may be partly attributed to differences between the magnitudes of news innovations. Relative contract mispricing in the NFL and NBA is contrasted in Figure 8.4, and it is clear that NBA contracts are relatively more price-efficient during active game play.

In the NFL market, the direction of news innovation also determines the magnitude of post-score mispricing. Following negative news innovations (those that drive price down), contracts are overpriced by \$9.47 on average, while after positive scores (those that drive price up), contracts are overpriced by \$4.97 on average. Evidently, there are relatively more willing buyers and fewer willing sellers after large unexpected price drops. This finding is consistent with a

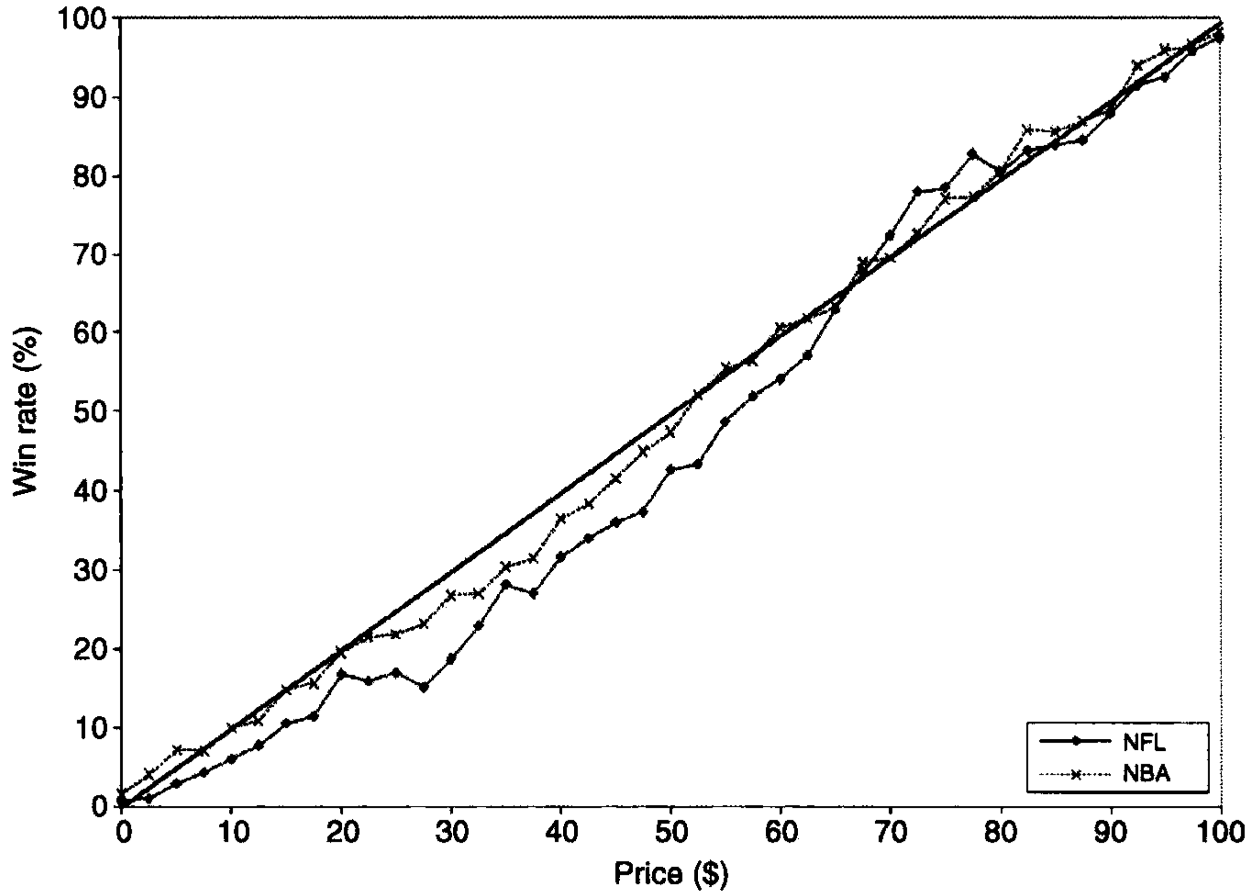


Figure 8.4 A plot generated from 216,564 (229,627) NFL (NBA) contract trades occurring after kickoff (tipoff). Each data point represents the observed mean win rate of contracts purchased at each price level (grouped into \$2.50 bands). A win is defined as expiry at \$100.

behavioral bias called the disposition effect, which is addressed below. One additional observation is that the magnitude of overpricing wanes as minutes pass following scores, so the market does react to mispricing, albeit at a slow pace.

2.6 Disposition effect

As mentioned earlier, prediction markets provide a superior laboratory to determine whether traditional equities investors act rationally. This is mainly due to problems associated with the joint hypothesis problem. In some instances, however, evidence of investor behavioral biases is so strong that the joint hypothesis problem does not pose a significant barrier to drawing reliable conclusions. One such instance is the disposition effect (i.e., the tendency of investors to show a preference towards realizing winning investments over losing investments). It has indeed been convincingly demonstrated (e.g., Weber and Camerer, 1998) that in traditional equities markets, investors prefer to sell winning stocks (those that have risen in price since purchase) and hold losing stocks (those that have dropped in price since purchase).

The simplest explanation for this phenomenon is that traders attempt to construct the most favorable self-image. To be a "winner," one must have bought low and sold high (thus, the incentive to sell winners). Additionally, one must avoid losses (thus, the incentive to hold losers, as selling at a loss is tantamount to admitting a mistake). Such behavior is clearly irrational, as the tax implications of such a strategy cause returns to be lower than otherwise (realizing a capital loss would result in a smaller loss, and realizing a capital gain would result in a smaller gain).

Hartzmark and Solomon (2009) find that prices are too low when a team gets ahead and too high when they get behind. This pattern is in harmony with what one would expect if the disposition effect that seems to be pervasive in equities markets also drives prices in prediction markets. It is no surprise that the disposition effect is present in the Tradesports NFL market, and potentially causes a significant portion of the observed mispricing.

2.7 Arbitrage opportunities

The price characteristics of prediction markets discussed thus far come from prior studies. I now present the results of original research, and describe its implications within the context of prior studies.

It is possible that when arbitrage operations are facilitated, some of the previously identified price biases will be mitigated or else disappear altogether. The significance of this idea is that if arbitrage opportunities do reduce price biases, then corporate officers should design internal prediction market contracts in such a way as to encourage arbitrage. Tradesports offered a set of conference and league championship contracts that may have done this. If the price efficiency of this set of contracts was significantly better than that of other contracts, then the format associated with the arbitrageable contracts would seem to be superior in generating unbiased forecasts.

The NFL is divided into two conferences – the American Football Conference (AFC) and the National Football Conference (NFC). There are 16 teams within each conference. Tradesports offered sets of NFL contracts whose values were determined by the likelihood that each team would win the conference and/or league championship (the Super Bowl). For example, from 2004 through 2006 there were 32 contracts (one for each team in the NFL) titled "NFLchamp. *TeamX*" where *TeamX* is one of the 32 teams. There were also 16 equivalent contracts each for the AFC and NFC championships.

If a trader believed that a particular team's league (conference) championship contract was overvalued, he or she could sell it and buy the remaining 31 (15) team contracts. This presents an arbitrage opportunity, and thus should help ensure that the prices of all related contracts are closer to true values than they otherwise would be in the absence of such an opportunity. If effective arbitrage operations occur, then the sum of contract prices across all teams should be no different than $100 (there is a 100 percent chance that one team will win the championship). If the sum is significantly different, then arbitrage operations are poorly functioning or non-existent.

In such an analysis, stale prices pose a significant problem; examination of the data reveals that not all conference and league championship contracts traded each week. To ensure that arbitrage operations were implementable, only weeks in which all team contracts traded are used in this analysis. The only times in which all team contracts were actively traded was in the week prior to the conference and league championship events; in 2004, 2005 and 2006, there are a total of 18 contracts that fit this criterion.

In Table 8.1, each contract is listed along with an arbitrage cohort indicator. For example, in 2004, the Raiders (Buccaneers) won the AFC (NFC) conference championship. The week before the Super Bowl, there was an arbitrageable opportunity (Arbitrage Opportunity 1). That is, the chance of the Buccaneers being champions must have been equal to one minus the probability of the Raiders being champions. Suppose the true probability of the Raiders winning the championship was 0.60, or 60 percent. Then it must be that the true probability of the Buccaneers winning the championship was $1-0.60=0.40$, or 40 percent.

If one or both of the contracts were mispriced, then a riskless profit-taking opportunity existed. Suppose the price of "NFLchamp.TeamA" is $65 and the price of "NFLchamp.TeamB" is $50. This is not out of the question, given the observed aggregate overpricing of NFL contracts on Tradesports. If this pricing were observed, then one could earn a riskless profit by simultaneously selling "NFLchamp.TeamA" and "NFLchamp.TeamB." The proceeds from the sale would be $65 + $50 = $115. No matter which team wins, the cost is guaranteed to be $100, resulting in a riskless net profit of $15. While this calculation ignores transaction costs, such costs are negligible compared to the massive overpricing of most NFL contracts.

Table 8.1 Arbitrage opportunities in the Tradesports NFL market, 2004–2006. A set of contracts is assumed to present an arbitrage opportunity when each of the contracts within the set trades within the week prior to the event

Contract number	*Arbitrage opportunity*	*Year*	*Title*	*Team*	*Bets*	*Mean price ($)*	*Sum ($)*	*Price adjustment ($)*	*Adjusted price ($)*	*Imputed sum ($)*
1	1	2004	NFL	Buccaneers	106	38.56	100.33	9.10	47.66	111.32
2	1	2004	NFL	Raiders	75	61.77		1.89	63.66	
3	2	2004	NFC	Buccaneers	31	16.06	101.10	2.60	18.66	121.44
4	2	2004	NFC	Eagles	15	34.27		9.10	43.37	
5	2	2004	AFC	Raiders	39	40.97		9.17	50.14	
6	2	2004	AFC	Titans	29	9.79		–0.53	9.26	
7	3	2005	NFL	Panthers	144	30.04	99.88	9.10	39.14	110.87
8	3	2005	NFL	Patriots	80	69.84		1.89	71.73	
9	4	2005	NFC	Colts	152	23.48	100.85	7.79	31.27	128.13
10	4	2005	NFC	Eagles	83	26.04		7.79	33.83	
11	4	2005	AFC	Panthers	105	11.93		2.60	14.53	
12	4	2005	AFC	Patriots	58	39.40		9.10	48.50	
13	5	2006	NFL	Eagles	1,205	28.94	100.04	7.79	36.73	113.03
14	5	2006	NFL	Patriots	663	71.10		5.20	76.30	
15	6	2006	NFC	Eagles	224	23.24	101.20	7.79	31.03	128.48
16	6	2006	NFC	Falcons	391	11.27		2.60	13.87	
17	6	2006	AFC	Patriots	403	38.99		9.10	48.09	
18	6	2006	AFC	Steelers	285	27.69		7.79	35.48	

Table 8.1 shows that the observed sum of "NFLchamp.Raiders" and "NFL-champ.Buccaneers" was $100.33. This combined price reflects a good deal of market efficiency. If no arbitrage opportunity had existed, and given the observed aggregate overpricing of assets within this price range, we would have expected the combined price of two assets to be $111.32. This number is obtained by adjusting upwards each contract's price by the observed overpricing within each price band. For instance, assets priced between $37.50 and $40.00 are overpriced by $9.10 on average, and assets priced between $60.00 and $62.50 are overpriced by $1.89 on average.[5] Thus, without the presence of Arbitrage Opportunity 1, we would have expected the sum of asset prices to be closer to ($38.56+$9.10)+($61.77+$1.89)=$111.32, but we instead observe that the combined price is only $100.33. The remaining arbitrage opportunities (2–6) consistently demonstrate that the sum of arbitrageable contract prices is much closer to $100 than would otherwise be expected.

This indicates that offering sets of contracts may be preferable to offering single contracts, because the presence of contract sets facilitates arbitrage opportunities. This has potentially important implications for corporations that use internal prediction markets to make business decisions. In this setting, the value of prediction markets is maximized when an unbiased forecast is produced. A set of internal corporate contracts might, for instance, take the form "Sales of Printer A will reach 1,000 units in May," "Sales of Printer B will reach 1,000 units in May," "Sales of neither Printer A nor sales of Printer B will reach 1,000 units in May" and "Sales of both Printer A and sales of Printer B will reach 1,000 units in May."

3 Summary and conclusion

During its period of active operations, Tradesports provided a valuable laboratory in which to test trader behavior and model tests of market efficiency. The majority of Tradesports contracts that were the focus of prior studies illustrate that considerable biases exist. The degree of pricing inefficiency is related to many factors, most notably contract price. Contracts priced below $65 are overvalued (buyers on average lose money), and those above $65 are undervalued (buyers on average gain). All contracts on average are overpriced, most likely because of a surplus of buyers and a shortage of sellers. This imbalance occurs because the vast majority of participants are accustomed to buying bets, and unaccustomed to short selling. The magnitude of these biases is so large that profit-taking opportunities exist.[6]

Greater trade volume does not alleviate this problem, as the added liquidity comes primarily from uninformed bettors, thus making wealth-constrained and risk-averse informed traders less able to push prices to equal rational values. The reverse favorite–longshot bias plays a significant role in mispricing as well, as all sides contracts are framed with respect to favorites, and because sports bettors are willing to pay a premium to bet on stronger teams.

Examination of bets placed after underlying games have begun also reveals that contracts are more overpriced immediately after negative events occur, especially when those events are critical to determining game outcome.[7] Finally, the disposition effect appears to affect trader behavior not just in traditional equities markets, but also in prediction markets. The net result is that prices are too low when teams get ahead and too high when they get behind.

While many of the observed price biases are large in magnitude, evidence suggests that the vast majority of these biases can be mitigated by offering contracts that are designed to avoid the reverse favorite–longshot bias and to facilitate arbitrage operations. For instance, constructing contracts so that bets are not framed with respect to a favorite, and also offering sets of contracts that are related to the same underlying event, would serve to reduce contract mispricing and thus generate more precise predictions.

Notes

1 Excluding transactions costs, if the probability of expiry at \$100 is 25 percent, then the expected contract value is (0.25)(\$100)+(1−0.25)(\$0)=\$25.
2 See Wolfers and Zitzewitz (2007) for a detailed explanation.
3 See Borghesi (2007) for the derivation of these numbers.
4 The traditional favorite–longshot suggests that bettors in pari-mutuel markets such as horse-racing over-bet on underdogs (longshots).
5 The overpricing estimates are extrapolated from data used to construct Figure 8.1.
6 Borghesi (2007) shows that selling contracts in the price band \$27.50±\$1.25 results in a mean return of 11.73 percent after transactions costs.
7 Borghesi (2007) shows that selling after negative scores even without conditioning on price results in a mean return of 16.12 percent after transactions costs.

References

Borghesi, Richard, 2007. "Price Biases in a Prediction Market: NFL Contracts on Tradesports," *The Journal of Prediction Markets* 1(3), 233–253.

Borghesi, Richard, 2009. "An Examination of Prediction Market Efficiency: NBA Contracts on Tradesports," *The Journal of Prediction Markets* 3(2), 65–77.

Borghesi, Richard, 2010 (forthcoming). "The Price Efficiency of Tradesports' Totals Contracts," *The Journal of Prediction Markets*, forthcoming.

Gruber, Martin, 1996. "Another Puzzle: The Growth in Actively Managed Mutual Funds," *The Journal of Finance* 51(3), 783–810.

Hartzmark, Samuel and David Solomon, 2009. "Efficiency and the Disposition Effect in NFL Prediction Markets." Working paper, University of Chicago.

O'Connor, Philip and Feng Zhou, 2008. "The Tradesports NFL Prediction Market: An Analysis of Market Efficiency, Transaction Costs, and Bettor Preferences," *The Journal of Prediction Markets* 2(1), 45–71.

Rubinstein, Mark, 2001. "Rational Markets: Yes or No? The Affirmative Case," *Financial Analysts Journal* 57(3), 15–29.

Schwert, G. William, 2001. "Anomalies and Market Efficiency." In G. Constantinides, M. Harris and R. Stulz (eds.), *Handbook of the Economics of Finance*, Amsterdam: North-Holland.

Tetlock, Paul, 2004. "How Efficient Are Information Markets? Evidence from an Online Exchange." Working paper, Columbia University.

Weber, Martin and Colin Camerer, 1998. "The Disposition Effect in Securities Trading: An Experimental Analysis," *Journal of Economic Behavior and Organization* 33(2), 167–184.

Wolfers, Justin and Eric Zitzewitz, 2006. "Five Open Questions About Prediction Markets." In Robert Hahn and Paul Tetlock (eds.), *Information Markets: A New Way of Making Decisions in the Public and Private Sectors*, Washington, DC: AEI-Brookings Press.

Wolfers, Justin and Eric Zitzewitz, 2007. "Interpreting Prediction Market Prices as Probabilities." Working paper, University of Pennsylvania and Stanford University.

9 The predictive ability of financial markets

Les Coleman

1 Introduction

A core assumption of market efficiency is that securities prices incorporate all price-sensitive information that is contained in historical prices. The usually tacit corollary is that price movements do not provide any guide to future prices, and hence markets have no predictive ability. However, an implicit challenge to this assumption is formal recognition that markets do have 'momentum' (Carhart, 1997): that is, price moves in one direction tend to continue. This is generally attributed to behavioural factors; namely, irrational biases in investor decisions ranging from herding and myopic trend-following to under – or over-appreciation of the significance of new information.

There is also, however, an argument that auto-correlation of security prices can occur with totally rational investors because prices of securities in markets with predictive ability will move towards a forecast price. Testing this argument is my research objective.

This chapter extends a decades old question (Cowles, 1933), and tests evidence and intuition that securities markets can predict future prices over the medium term.

2 The intuition behind markets' predictive ability

It is simple to explain why markets *should* have predictive ability. Take the standard assumption in finance that security prices equal the present value of expected future cash flows discounted at a risk-adjusted rate of return. That is, looking n periods ahead, the price of a security at time t is:

$$\text{Price}_t = \sum_{x=1}^{n} \frac{\text{CashFlow}_{t+x}}{(1+k_e)^x} + \frac{\text{Price}_{t+n}}{(1+k_e)^n} \tag{9.1}$$

In words, a security today is priced at its expected future price plus the value of cash flows expected to be received in the interim, all discounted at a risk-adjusted rate k_e.

In a multi-period model, the price of a security will move towards its expected future price as shown below.

$$\text{Price}_{t+1} = \sum_{x=2}^{n} \frac{\text{CashFlow}_{t+x}}{(1+k_e)^x} + \frac{\text{Price}_{t+n}}{(1+k_e)^{n-1}} \tag{9.2}$$

Subtracting equation (9.1) from (9.2):

$$\text{Return}_{t,t+1} = \text{Price}_{t+1} - \text{Price}_t$$

$$= \sum_{x=2}^{n} \frac{\text{CashFlow}_{t+x}}{(1+k_e)^x} + \frac{\text{Price}_{t+n}}{(1+k_e)^{n-1}} - \sum_{x=1}^{n} \frac{\text{CashFlow}_{t+x}}{(1+k_e)^x} - \frac{\text{Price}_{t+n}}{(1+k_e)^n}$$

$$= \left(\frac{\text{Price}_{t+n}}{(1+k_e)^{n-1}}\right) - \left(\frac{\text{Price}_{t+n}}{(1+k_e)^n}\right) + \left(\sum_{x=2}^{n} \frac{\text{CashFlow}_{t+x}}{(1+k_e)^x} - \sum_{x=1}^{n} \frac{\text{CashFlow}_{t+x}}{(1+k_e)^x}\right)$$

$$= \left(\frac{\text{Price}_{t+n}}{(1+k_e)^{n-1}} - \frac{\text{Price}_{t+n}}{(1+k_e)(1+k_e)^{n-1}}\right) +$$

$$\left(\sum_{x=2}^{n} \frac{\text{CashFlow}_{t+x}}{(1+k_e)^x} - \sum_{x=1} \frac{\text{CashFlow}_{t+x}}{(1+k_e)^x} - \sum_{x=2}^{n} \frac{\text{CashFlow}_{t+x}}{(1+k_e)^x}\right) \tag{9.3}$$

Further

$$\text{Return}_{t+1,t+2} = \text{Price}_{t+2} - \text{Price}_{t-1}$$

$$= \frac{k_e \cdot \text{Price}_{t+n}}{(1+k_e)^{n-1}} - \frac{\text{CashFlow}_{t+2}}{(1+k_e)^2} \tag{9.4}$$

A complementary perspective uses the Gordon (1962) dividend discount model where dividends, D, grow indefinitely at a constant rate g:

$$P_t = \frac{D_{t+1}}{k_e - g}; P_{t+1} = \frac{D_{t+2}}{k_e - g} = \frac{(1+g)*(D_{t+1})}{k_e - g}; P_{t+2} = \frac{D_{t+3}}{k_e - g} = \frac{(1+g)^2 \times (D_{t+1})}{k_e - g}$$

etc.

Thus

$$\text{Return}_{t,t+1} = \frac{D_{t+1}(1+g) - D_{t+1}}{k_e - g} = \frac{g * D_{t+1}}{k_e - g} \tag{9.5}$$

$$\text{Return}_{t+1,t+2} = \frac{D_{t+1}(1+g)^2 - D_{t+1}}{k_e - g} = \frac{g^2 * D_{t+1}}{k_e - g} \tag{9.6}$$

By unpacking common finance assumptions, it is clear that any security's price will move in line with the market's expectation of risk and returns. In other words, security returns are determined by ex ante expectations of future cash flows, price and risk. Moreover, if cash flows are regular, then returns will be auto-correlated and will trend. In the case of constantly growing cash flows (as in the Gordon model), prices will move geometrically towards their expected future value. As shown above, returns reflect the market's estimation of the future price and hence are a guide to markets' predictive abilities.

At the extreme where there is a single cash flow, the security will be priced at below the value of the cash flow and then gradually approach it. Thus, the security's past direction and rate of change will – in the absence of changed expectations about the security – persist. This is a fundamental support for expectations of the long-term increase in security prices and hence positive investment returns.

So why do prices of securities, especially equities, fall? At the time of writing, for instance, the MSCI World Index and the US S&P 500 have seen no increase in the past decade; so most investments during the period would have lost money. Equations (9.3) and (9.4) show that security price changes during any period are driven by the cash flow expected during the period. Thus a security price will fall when a large cash flow is received (which is why stock prices drop on their ex-dividend date). The only other explanation for a share price decline is that the expected return or risk is revised, which is *prima facie* evidence that the market has not been able to predict future cash flows and risk.

Thus, simple accounting calculus shows that – whether security prices are rising or falling – if markets have predictive ability then the direction and rate of change in security price should persist over the market's effective forecast horizon. As a corollary, a constant return vector (direction and speed of movement) as evidenced by momentum or auto-correlation of securities prices is evidence that investors have predictive ability, and vice versa.

Even in markets with predictive power, it is probable that this ability drops off with forecast horizon. With the emergence of imperfectly predictable events or pieces of new information, markets re-assess the outlook and eventually make significant changes in the return vector, and a regime shift occurs (Hamilton, 1989). The predictive ability of any market is a function of the duration of its regimes.

This argument is not represented in the literature, which is typically built around the concept of market efficiency such that 'prices on average adjust quickly to firm-specific information' (Fama, 1991: 1062), and so 'the current equilibrium price fully reflects all available information and price adjusts instantaneously to new information' (Park and Irwin, 2007: 805). The unstated assumption is that available information implies the same real security price across time. Alternatively, today's security price (possibly adjusted for holding costs) is the best estimate of its future price.

Conversely, the discussion above shows that even in the absence of change in available information about a security, the effluxion of time changes the discounted value of this information and so predictably affects security prices. The critical point behind my intuition is that even if markets are fully efficient and accurately value a security, today's price is not that at which the security will trade in the future.

Turning to the evidence on markets' predictive ability, if none is demonstrated there are two probable explanations: investors do not have any skill in forecasting future returns and risk; or factors other than expected returns and risk dominate investor judgements. Consider each in turn.

The literature on forecasting is generally not kind to experts in any discipline. Camerer (1995) identified 100 studies that have been carried out to determine the accuracy of experts in forecasting outcomes from processes supported by observable data such as test scores, medical examinations and financial ratios. The consistent finding is that simple regressions using historical data outpredict the experts. He concludes (Camerer, 1995: 611): 'The *only* documented exceptions to the general conclusions that models outpredict experts are a few kinds of esoteric medical diagnosis [emphasis in original].' A similar picture is painted by Yates (1990), who found that physicians, psychologists and stock analysts could do no better than simple naive strategies; only weather forecasters consistently display skill. Bolger and Wright (1994) identified 40 studies of expertise, but found only six groups of experts that evidenced good judgement, including race-track tipsters.

In macroeconomic forecasting, Atkeson and Ohanian (2001) found that expert inflation forecasts are little better than chance. Loungani (2000) concluded that private sector GDP forecasts have significant errors, with an upward bias of about 1 per cent for a year ahead, largely due to excessive optimism about developing economies. Schuh (2001: 36) evaluated the performance of private US forecasters between 1969 and 2000, and concluded that 'there is ample evidence that average macroeconomic forecasts are not efficient'. Even though there are plentiful incentives for expertise amidst the complexity of markets, evaluations of experts' performance in the economics and finance literature conclude that a coin remains an effective tool (Cochrane, 1999).

While a comprehensive evaluation of forecasting ability is beyond the scope of this chapter, there is strong evidence that few experts in any discipline add value. One reason is that experts are not able to identify and leverage cues that predict outcomes. This is termed the *missing variables problem*, and refers to the fact that data on significant variables are either not available to the analysts or are not included in their forecast. The second reason for forecast errors is that the rate of flow of new information is high and swamps the predictive power of previous information.

The other explanation of investors' inability to forecast security returns and risk comes from evidence that, even though security prices are assumed to reflect rational expectations, irrational factors have extensive influence. These include biases in investor decisions, herding or myopic trend-following by noise traders, and over or under-reaction to news by investors through misjudging its significance (Ricciardi, 2008).

Looking ahead, the next section reviews the literature on the predictive ability of conventional financial markets. I then use simple tests to examine hypotheses developed from the assumption that markets such as the US S&P 500 have predictive ability, and close with a discussion.

3 Literature on markets' predictive ability

A number of studies have examined whether models of security prices have predictive power. One stream of work has looked at individuals' skill in forecasting, such as the finding by Barber *et al.* (2001) that a strategy of buying stocks recommended by security analysts and short-selling those that are least favourably recommended is able to outperform the market by around 8 per cent per year. Another stream looks at quantitative asset pricing models, and generally finds that they demonstrate some predictive ability (Simin, 2008). More specifically, there is a positive, statistically significant relationship between US stock returns and the dividend: price ratio (Cochrane, 2008), at least for up to a few months (Ang and Bekaert, 2006). As a result, leading valuation models are developed on the basis of the link between stock yields and interest rates (Campbell and Vuolteenaho, 2004).

Of more relevance to the predictive power of security prices, research has identified momentum as a factor in explaining security returns. Carhart (1997), for instance, employs the three Fama and French (1996) factors (market return, size and book-to-market ratio) and adds a fourth factor termed 'momentum in stock returns'. The last is a market measure, which at time t reflects the difference in returns during the previous period, that is, from $t-1$ to t, between portfolios of securities that had the highest and lowest returns in the preceding n periods, that is from $t-1$ n to $t-1$. Momentum is equivalent to trend persistence by the market, and strategies based on it can deliver significant positive returns (Jegadeesh and Titman, 2001).

Obvious applications of momentum's predictive power include numerous technical analysis techniques. These are based on the intuition that prices trend in accordance with investor attitudes, and trade these trends using indicators, including those developed from market data, particularly prices (Pring, 2002). A trend can arise because markets have predictive ability; or else because security prices are determined by supply and demand factors and hence trend until there is a realignment of the supply–demand balance (Pinches, 1970).

The most popular technical strategy relies on moving averages, which recognise that stock prices can be volatile and dampen short-term fluctuations to provide a smooth trend that gives a better indication of momentum-based direction. The simplest application averages the (say) ten most recent prices, and identifies a change in trend when the price moves up or down through its moving average, or two moving averages with different numbers of prices which cross over one another (Pring, 2002).

In a metastudy of technical trading strategies, Park and Irwin (2007) review the technique's history, and conclude that about half of the 96 studies published between 1988 and 2004 showed that technical analysis of equity, foreign exchange and futures markets techniques was profitable, although probably only until the mid-to-late 1990s.

Tests of technical strategies that are of most relevance to this chapter are those conducted in deep equity markets, especially using the major US indices. This choice is deliberate and recognises that success of technical indicators is often attributed to market anomalies such as systematic under-reaction to new information, trend chasing or market over-reaction (Park and Irwin, 2007). In addition, mis-pricing can arise from market frictions and lower trading volumes that slow price discovery. An indicator of the significance of such confounding factors is that simple technical trading rules provide a higher return in emerging markets than developed markets (Fifield *et al.*, 2008). While anomalies can occur in all markets, they should be minimised by examining deep markets that are well served by analysts and are dominated by informed investors.

An early study by James (1968) set the standard for testing markets' predictive power, which is that the accuracy of a forecast using historical price data is superior to that of a forecast using no data. Put differently, if markets work randomly, then no trading strategy can deliver a profit relative to return from the market: changes in security prices should be independent of price history.

Table 9.1 reports results of a variety of subsequent studies, and suggests that the literature is, at best, ambiguous on whether historical prices convey valuable information. A major reason is that most analyses of the predictive power of prices are questionable because they use them in conjunction with ex post trading rules or models. Thus, moving averages – which are currently seen as the most popular trading technique – are back-tested, as are dividend: price ratios and the like. The net is that trading rules and techniques that have become popular (presumably because they have worked in practice) are evaluated using the same historical data that led to their popularity. Even though analysts may claim that tests are conducted out of sample, their results are conditional on choosing their test techniques from popular models, and in reality are in-sample tests that suffer from bias in model selection.

One approach to guard against this data snooping is to test a much wider pool of techniques than just those that have survived. Sullivan *et al.* (1999) attempted to test performance of a large universe of technical trading rules. Another approach, which is taken in the following section, is to eschew any model, and simply test the hypothesis that returns are auto-correlated, which (as discussed in Section 1) would evidence markets' predictive power.

In summary, use of technical analysis techniques is common amongst practitioners, particularly for timing trades identified from more sophisticated value-based analysis. Baker and Wurgler (2002), for example, provide evidence of the last by showing that managers time the equity market in decisions on issuing equity. This is in marked contrast to the traditional view in academic literature that technical analysis using historical prices cannot add value.

Table 9.1 Sample of published studies of technical analysis in US markets

Period covered	*Securities analysed*	*Findings*	*Author (date)*
1926–1960	CRSP stocks	Tested portfolios compiled using eight decision rules with monthly data. None outperformed buy-and-hold	James (1968)
1897–1986	Dow Jones Index	Moving averages using daily data produce an incremental return of 12 per cent per annum, with better results from buy signals	Brock *et al.* (1992)
1926–1991	Daily DJIA	Found excess annual return of over 3 per cent	Bessembinder and Chan (1998)
1897–1996			
1984–1996	DJIA S&P 500 futures	Tested 8,000 variants of technical analysis and found short-term forecasting ability	Sullivan *et al.* (1999)
1897–2000	Daily DJIA	Re-evaluated the conclusion of Brock *et al.* (1992) using a different sample and found poor results	Ready (2002)
1928–1995	Daily S&P 500	Technical trading does not outperform buy-and-hold	Allen and Karjalainen (1999)
1962–1996	US stocks	Tested chart patterns using algorithms and found several offer good results, especially in NASDAQ stocks	Lo *et al.* (2000)
1996	Intraday spot foreign exchange rates for four currencies	Analyses of foreign exchange markets using two methodologies founds no evidence of excess returns	Neely and Weller (2003)

4 Simple tests of markets' predictive ability

A problem with tests of market predictive power is that they use models with built-in assumptions, such as moving averages, or else employ trading rules developed from market data. This raises the *joint test problem*, where it is not clear if results showing price persistence or not are driven by historical prices or by the hypothesised model.

The approach taken here is a-theoretical and uses two simple tests. Data comprise returns from the S&P 500 Index over the period 2005–2009, which is the most recent full cycle in the US equity market and eliminates biases from analyses during long unbroken trends in markets that have been shown to significantly distort results (e.g. Bennett and Sias, 2006). Returns are chosen, rather than prices, because prices are known to be non-stationary, and auto-correlation will arise from the time-series properties of the data.

The first test examines whether markets have a 'memory', which is a precondition to predictive ability where the accuracy of a forecast that incorporates historical data is greater than the accuracy of a naive forecast without any historical data. This is analogous to Granger causality, where historical returns are said to cause future returns if their lagged values help predict current future return when incorporated in an equation with current return (Gujarati, 1995); that is, Historical returns Granger-cause Future return if $\Pr(\text{Return}_t \mid \text{Return}_{t-1}, \text{Return}_{t-2}, \text{Return}_{t-3}, \ldots) > \Pr(\text{Return}_t \mid \text{Return}_{t-1})$. Granger causality implies that past plus current values explain future values better than current value alone: thus the test is not one of strict cause-and-effect, but rather of enhanced prediction. The null hypothesis is that historical returns do not Granger-cause return, and is rejected if the coefficients on historical returns are significant. The regression to be examined is:

$$\text{Return}_t = \alpha + \sum_{k=1}^{n} \beta_k \cdot \text{Return}_{t-k} + \varepsilon \tag{9.7}$$

Table 9.2 reports results using daily, weekly and monthly values of S&P 500 returns during 2005–2009, with lags of up to 12 periods. This shows that daily returns from the previous one and three days have significant ($p<0.01$) positive relationship to current return. There is, however, no evidence of statistically meaningful ($p<0.05$) predictive ability for daily returns beyond three days. Importantly, weekly and monthly returns show no predictive ability.

A second test looks at the duration of trends. A trend is defined as a continuous sequence of price changes above a minimum cut-off. The intuition is that markets with predictive power will establish trends that are more likely to continue than to collapse. Table 9.3 reports results using daily, weekly and monthly returns with lags of two, three and five periods and cut-offs for minimum returns of 0, 0.1 and 0.25 per cent.

The results show no evidence of predictive ability in market uptrends. In terms of downtrends, daily data show significant mean reversion after two or

Table 9.2 Causality of lagged prices

Data frequency	*Selected lag lengths (k)*						*Adj R^2*
	1	*2*	*3*	*5*	*8*	*12*	
Daily	0.852***	0.041	0.152***	0.017	0.065*	0.028	0.994
Weekly	−0.068	0.013	−0.025	−0.004	−0.042	−0.014	0.252
Monthly	−0.089	−0.010	−0.024	0.026	−0.038	0.077	0.572

Notes

This table reports results from regression of: $\text{Return}_t = \alpha + \sum_{k=1}^{n} \beta_k \cdot \text{Return}_{t-k} + \varepsilon$. Data in each case is return from the S&P 500 Index during 2005–2009. Values of β are reported for selected lags (i.e. values of k) of between 1 and 12. Significance of coefficients is indicated by: * $p<0.1$; ** $p<0.05$; and *** $p<0.01$.

Table 9.3 Predictive ability of trends

Data frequency	*Uptrend (periods)*			*Downtrend (periods)*		
	2	*3*	*5*	*2*	*3*	*5*
Panel A: minimum move per period > 0 per cent						
Daily	49.3	46.4	42.9	38.5***	31.8***	44.4
Weekly	47.8	51.5	33.3	42.6	35.0	50.0
Monthly	40.0	50.0	0.0	71.4*	66.7	71.4
Panel B: minimum move per period > 0.1 per cent						
Daily	49.4	47.6	38.9	37.2***	33.3**	50.0
Weekly	46.7	58.3	40.0	41.5	29.4*	n/a
Monthly	40.0	50.0	0.0	75.0*	66.7	100.0
Panel C: minimum move per period > 0.25 per cent						
Daily	49.4	46.9	40.0	36.7***	33.3**	50.0
Weekly	51.1	66.7	50.0	43.6	29.4*	n/a
Monthly	40.0	50.0	0	75.0	66.7	100.0*

Notes

This table shows the proportion of changes in any period that are in the same direction as a prior continuous sequence of price changes greater than the minimum move and in the same direction. That is, it shows persistence of trends that have been established over prior periods. Significance versus a random binomial outcome (that is either up or down) is indicated by: * $p<0.1$; ** $p<0.05$; and *** $p<0.01$.

three days of falls. That is, only about one-third of two- or three-day downtrends are followed by another fall in price. Otherwise, there is no statistically meaningful ($p<0.05$) evidence that trends have predictive ability. Beyond two or three days, an established up or down price trend is equally likely to be followed by a positive or negative return, no matter which of a range of criteria are used to define the trend.

Results in Tables 9.2 and 9.3 are consistent. Weekly and monthly data show nothing that is statistically meaningful ($p<0.05$). Daily data show significant evidence of trending and mean reversion of downtrends over two or three days, but little else.

5 Summary and conclusions

The hypothesis underpinning my analysis is that there should be medium-term trends in the returns of securities in markets that price securities on the basis of expected returns and risk and that have predictive ability. This hypothesis is tested without imposing any model or structure, which differs from traditional evaluations of the forecasting ability of historical prices, which test commonly used rules and thus inevitably suffer from the joint test problem and survivorship bias.

The only statistically meaningful result of my analysis using daily, weekly and monthly data for the major US equity index through its last cycle is that daily returns persist for two or three days and that declines lasting two or three days are followed by mean reversion. In short, the US S&P 500 Index has nothing more than trivial, short-term forecasting ability. The horizon of the market's accurate forecasts is well inside the frequency of cash flows, and so does not arise through any ability to predict cash flows or risk.

This has a number of important implications.

The first is to question the core finance assumption that securities are priced on the basis of expected return and risk. If this were true, then (as outlined in Section 1) markets should trend towards expected prices over periods of months or more. Simple analyses here show that the accurate forecasting horizon of investor expectations is no more than a few days. This means the volume of unexpected new price-sensitive information almost continuously swamps the stock of expected information and markets are forced every few days to update expectations of returns and risk. In short, current, revealed security prices are poor estimates of future prices and hence of securities' intrinsic value.

The second implication is that historical prices have no predictive ability, which means that technical rules derived from them cannot have economic significance. Thus, any 'successful' trading system must arise from chance, temporary patterns or other transient market influences, and so will not succeed over the longer term. This, in fact, is evidenced by comparative studies of trading systems across different periods, which show that commonly used models (which have presumably become popular because of their success) are rarely as successful in subsequent periods. The predictive weakness of technical analysis does not arise from market efficiency – in which historical price data are impounded in current prices – but from the lack of forecasting ability of investors. This, too, matches evidence that few investors display predictive skill.

It is important to note that, while these conclusions deny any predictive ability to historical security returns, they are silent on the predictive ability of other security attributes such as dividend: price ratio that underpin asset valuation models. My results relate solely to weak-form market efficiency and say nothing about semi-strong efficiency, much less strong-form efficiency (Fama, 1970). Skill and inside knowledge, then, *may* be useful.

What are the implications of my results for prediction markets? They only pertain to the predictive ability of historical security prices in conventional markets and find they have none beyond the very short-term. The implication of this, though, is that investors in aggregate have no ability to predict future returns and risks. If they did possess such an ability, security prices should move inexorably towards expected future prices, and so historical prices should be related to future prices. Although investors generally lack predictive ability, many studies have identified subsets of analysts with skill who can outperform chance. It is probably the skill of this latter group with predictive ability that explains any outperformance in prediction markets.

References

Allen, F. and Karjalainen, R. (1999). Using genetic algorithms to find technical trading rules. *Journal of Financial Economics* 51 (2): 245–271.

Ang, A. and Bekaert, G. (2006). Stock return predictability: is it there? *Review of Financial Studies* 20 (3): 651–707.

Atkeson, A. and Ohanian, L.E. (2001). Are Phillips curves useful for forecasting inflation? *Federal Reserve Bank of Minneapolis Quarterly Review* 25: 2–11.

Baker, M. and Wurgler, J. (2002). Market timing and capital structure. *The Journal of Finance* 57 (1): 1–32.

Barber, B., Lehavy, R., McNichols, M. and Trueman, B. (2001). Can investors profit from the prophets? Security analyst recommendations and stock returns. *The Journal of Finance* 56 (2): 531–563.

Bennett, J.A. and Sias, R.W. (2006). Why company-specific risk changes over time. *Financial Analysts Journal* 62 (5): 89–100.

Bessembinder, H. and Chan, K. (1998). Market efficiency and the returns to technical analysis. *Financial Management* 27 (2): 5–17.

Bolger, F.A. and Wright, G. (1994). Assessing the quality of expert judgement. *Decision Support Systems* 11 (1): 1–24.

Brock, W., Lakonishok, J. and LeBaron, B. (1992). Simple technical trading rules and the stochastic properties of stock returns. *Journal of Finance* 47 (6): 1731–1764.

Camerer, C. (1995). Individual decision making. In J.H. Kagel and A.E. Roth (eds), *The Handbook of Experimental Economics*. Princeton University Press, Princeton, NJ.

Campbell, J.Y. and Vuolteenaho, T. (2004). Inflation illusion and stock prices. *The American Economic Review* 94 (2): 19–23.

Carhart, M.M. (1997). On persistence in mutual fund performance. *The Journal of Finance* 52 (1): 57–82.

Cochrane, J.H. (1999). New facts in finance. *Economic Perspectives* 23 (3): 36–58.

Cochrane, J.N. (2008). The dog that did not bark. *Review of Financial Studies* 21 (4): 1533–1575.

Cowles, A. (1933). Can stock market forecasters forecast? *Econometrica* 1: 309–324.

Fama, E.F. (1970). Efficient capital markets: a review of theory and empirical work. *The Journal of Finance* 25 (2): 383–417.

Fama, E.F. (1991). Efficient capital markets: II. *The Journal of Finance* 46 (5): 1575–1617.

Fama, E.F. and French, K.R. (1996). Multifactor explanations of asset pricing anomalies. *The Journal of Finance* 51 (1): 55–84.

Fifield, S.G.M., Power, D.M. and Knipe, D.G.S. (2008). The performance of moving average rules in emerging stock markets. *Applied Financial Economics* 18: 1515–1532.

Gordon, M.Y. (1962). *The Investment, Financing, and Valuation of the Corporation*. Irwin, Homewood, IL.

Gujarati, D.N. (1995). *Basic Econometrics*. McGraw-Hill, New York.

Hamilton, J.D. (1989). A new approach to the economic analysis of non-stationary time series and the business cycle. *Econometrica* 57: 357–384.

James, F.E.J. (1968). Monthly moving averages: an effective investment tool? *The Journal of Financial and Quantitative Analysis* 3 (3): 315–326.

Jegadeesh, N. and Titman, S. (2001). Profitability of momentum strategies: an evaluation of alternative explanations. *The Journal of Finance* 56: 699–720.

Lo, A.W., Mamaysky, H. and Wang, J. (2000). Foundations of technical analysis: computational algorithms, statistical inference, and empirical implementation. *The Journal of Finance* 55 (4): 1705–1765.

Loungani, P. (2000). *How Accurate are Private Sector Forecasts? Cross-country Evidence from Consensus Forecasts of Output Growth*. Working Paper WP/00/77. International Monetary Fund, Washington, DC.

Neely, C.J. and Weller, P.A. (2003). Intraday technical trading in the foreign exchange market. *Journal of International Money and Finance* 22 (2): 223–237.

Park, C.-H. and Irwin, S.H. (2007). What do we know about the profitability of technical analysis? *Journal of Economic Surveys* 21 (4): 786–826.

Pinches, G.E. (1970). The random walk hypothesis and technical analysis. *Financial Analysts Journal* 26 (2): 104–110.

Pring, M.J. (2002). *Technical Analysis Explained*. McGraw-Hill, New York.

Ready, M.J. (2002). Profits from technical trading rules. *Financial Management* 31 (3): 43–61.

Ricciardi, V. (2008). The psychology of risk: the behavioral finance perspective. In F.J. Fabozzi (ed.), *Handbook of Finance: Volume 2: Investment Management and Financial Management*. John Wiley & Sons, New York.

Schuh, S. (2001). An evaluation of recent macroeconomic forecast errors. *New England Economic Review*, January/February: 5–35.

Simin, T. (2008). The poor predictive performance of asset pricing models. *Journal of Financial and Quantitative Analysis* 43 (2): 355–380.

Sullivan, R., Timmermann, A. and White, H. (1999). Data-snooping, technical trading rule performance, and the bootstrap. *The Journal of Finance* 54 (5): 1647–1691.

Yates, J.F. (1990). *Judgment and decision making*. Prentice Hall, Upper Saddle River, NJ.

10 The ability of markets to predict conditional probabilities

Evidence from the US presidential campaign

Lionel Page

1 Introduction: from prediction markets to decision markets?

Prediction markets give an estimation of the probability of an event *A*. Such a prediction is very useful to decision-makers in all the situations where the probability of event *A* matters in choosing between two options. For instance, let $S \in \{0, 1\}$ be the set of possible states of nature, $D \in \{0, 1\}$ a set of possible decisions and *V* a utility function defined on *S* and *D*, knowing the probability of the event $S=1$ will be useful in solving problem:

$$\max_{D} \left\{ \sum_{S} V(D,S)P(S) \right\}. \tag{10.1}$$

However, in most situations, the decision-maker expects his decision to have an effect on the state of the world, and she needs to solve the following problem, where conditional probabilities replace simple probabilities:

$$\max_{D} \left\{ \sum_{S} V(D,S)P(S \mid D) \right\}. \tag{10.2}$$

Typically, a decision-maker needs an answer to the question: 'What happens if one decision is taken instead of another one?'. The estimation of conditional probabilities would therefore be a major qualitative advance relative to the estimation of simple probabilities. A policy-maker could, for instance, ask:

- What is the best policy to reduce crime? Tougher prison sentences? More police?
- What is the best policy to limit/reduce global warming? Increase growth? Reduce inequality?

In turn, an elector could choose the best candidate, knowing the probability of success of each candidate on a set of key issues (criminality, growth, unemployment, etc.). It is therefore not surprising that the estimation of conditional probabilities by prediction markets has been one of the most important hopes in the field in the last ten years. Some have called these new estimation tools 'decision markets' (Hanson, 1999, 2007) or 'contingent markets' (Wolfers and Zitzewitz, 2004; Henderson *et al.*, 2008).

The simplest way to obtain conditional probabilities is to use the available design of prediction markets. Using a simple Bayes rule, one can use two prediction markets to estimate a conditional probability. Suppose we have a market on '*B* will happen' and a market on '*A* and *B* will happen jointly', then the conditional probability of *A* happening if *B* happens is:

$$P(A \mid B) = P(A \cap B) / P(B). \tag{10.3}$$

Such a way of calculating conditional probabilities has been used by Berg and Rietz (2003), Wolfers and Zitzewitz (2004) and recently by Mankiw (2008).

This said, no empirical work has yet been done to assess the quality of these estimators of conditional probabilities. Rightly, it is assumed that if prediction markets give a good estimate of $P(A \cap B)$ and $P(B)$, then the ratio must by definition be a good estimate of $P(A|B)$. It has, however, been shown that prediction market prices are subject to some biases. The question of the calibration of the estimation of conditional probabilities is therefore an empirical question that has still to be answered.

I present here the first study of the quality of the estimation of conditional probabilities by prediction markets using the unique opportunity given by the existence of the markets for the primaries in the race for the nomination to the US presidential election in 2008. The race for the nomination presents an ideal situation since we have at our disposal the probability of winning the nomination process for each candidate P_{prim} and the probability of this candidate winning the presidential elections P_{pres}. As a consequence, it is possible to estimate the probability of winning the election *if* nominated:

$$\frac{P_{pres}}{P_{prim}}. \tag{10.4}$$

This estimation is certainly of primary interest for voters in each party. The markets on the primaries and presidential campaign also present the advantage of being very liquid and of spanning a long period of time, which makes them a perfect pool to study the estimation of conditional probabilities by prediction markets.

The rest of the chapter develops as follow. Section 2 stresses that prediction markets often face a major challenge in estimating conditional probabilities as they are characterised by low liquidity when their duration is very long. Section 3 shows that the estimation of conditional probabilities by prediction markets suffers from a large variance if the probability of the joint event is relatively small. Section 4 lists a set of potential biases and yields some evidence of these biases using the data on the primary race. Section 5 questions the real meaning of these estimations and Section 6 concludes.

2 Liquidity

To predict conditional probabilities, prediction markets must be attractive enough for traders to create liquid markets. However, two problems exist here. First, the decision-maker may have to choose between actions which have mid- or long-term effects. Therefore 'success', where the probability for each decision is estimated, must be measured over a horizon of several years. If a policy-maker wants to choose the best policy to increase growth or reduce crime, it is necessary to wait more than a few months to see if and how the rate of growth or crime changes. If the CEO of a company hesitates between two choices of technical innovations for the design of a product, the final success or failure of the product will only be observed after the product has been designed and produced. If a policy-maker is unsure as to the best way to reduce global warming, the effect of any policy will be observed only in several years.

In these examples, the decision-maker may need prediction markets that predict events over a reasonable time horizon. For instance, in the extreme case of global warming, a market would have to predict something like 'Policy A is implemented and a positive effect on global warming is observed ten years later'.

This need to predict events taking place over a reasonable time horizon is a problem for prediction markets. As identified by Page and Clemen (2008), the time preferences of traders make long-term prediction markets unattractive. The reason is simple – if a trader has a preference for the present, she will dislike investing in a market where her money will be unavailable for a long time. As a consequence, in a long-term prediction market, a risk-neutral trader will not always buy a contract if her belief of the probability of the event is higher than the observed price. She will need a premium to exist between the observed price and her belief in order to buy. This premium compensates for the disutility of freezing an amount of money for a long period of time. If, for instance, a trader would buy a contract expiring tomorrow (whose real probability she believes is 61) at 60 today, she would need her beliefs to be at least 65 if the contract ends in several months.

The effect of the traders' preference for the present is to reduce the support of beliefs for which there can be some trade in the market. While the premium increases, the number of traders having beliefs compatible with a trade on the market decreases.

Figure 10.1 shows the evolution of the volume of trade on the presidential markets as a function of the number of days left to expiration. While the markets are very liquid towards the time of resolution of the markets' uncertainty, there are hardly any trades two years before. The presidential markets generate great interest among traders and such a feature would be even stronger for markets about less popular topics.

This lack of liquidity for long-term markets is problematic for any estimation of conditional probabilities that requires the estimation of some success in the long-term. Estimation of conditional probabilities may, in this case, face a practical impossibility because no trader is willing to invest in these markets. To solve this problem, Page and Clemen (2008) suggest a solution that has already

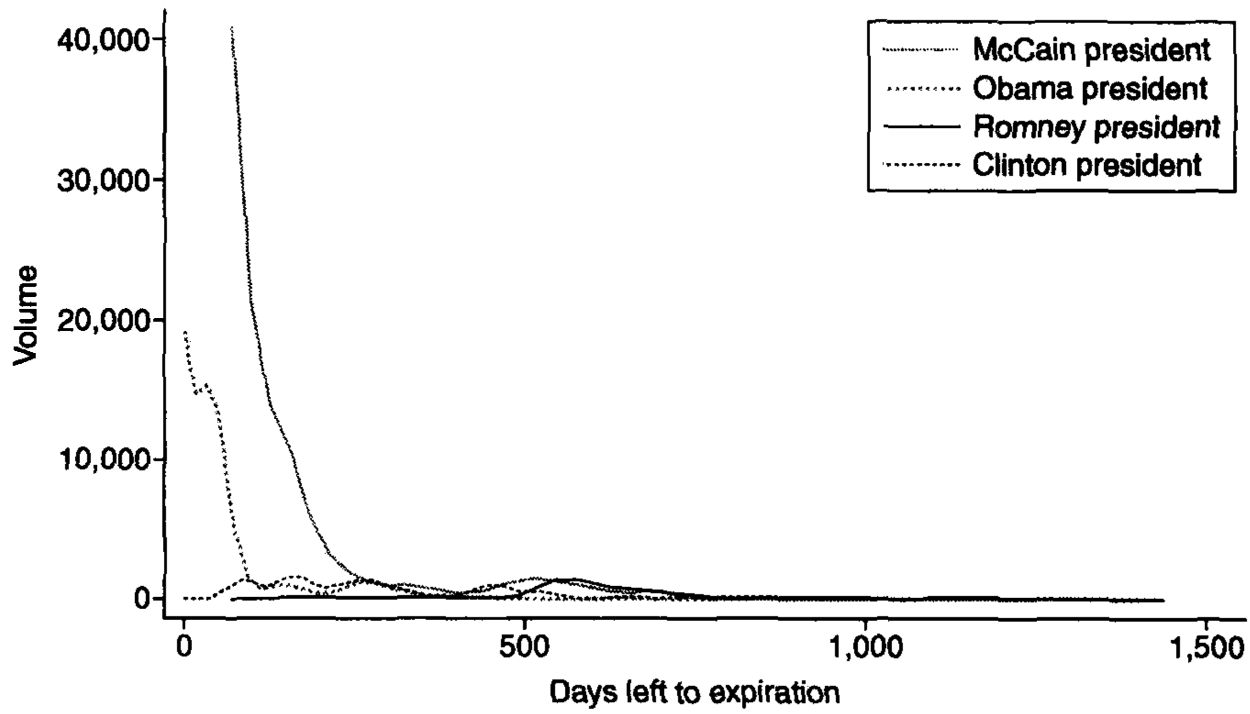


Figure 10.1 Low volume in the long-term.

been raised by the proponents of decision markets: to create prediction markets whose contracts have an increasing value with time (similar to a rate of interest). However, such a design requires the betting market operator to be able to place the amount of money bet in the bank or on the financial market to create interest. This may conflict with betting market regulations, which in some cases constrain the operators to keep the amount invested.

3 Variance

In the situation where the prediction markets are liquid and where it is possible to estimate conditional probabilities, new questions arise. In particular, the estimation of conditional probability uses a ratio of prediction market prices, which raises questions about the properties of such an object.

The first thing to consider is that a ratio will typically be more sensible to variations when the values of its elements are smaller. In the case of the estimation of the conditional probability of winning the presidency if nominated, one can see that the derivative of the conditional probability (CP) is higher for lower values of P_{pres} and P_{nom}:

$$\frac{\partial \frac{P_{pres}}{P_{nom}}}{\partial P_{nom}} = \frac{P_{pres}}{P_{nom}^2}, \tag{10.5}$$

$$\frac{\partial \frac{P_{pres}}{P_{nom}}}{\partial P_{pres}} = \frac{P_{pres}}{P_{nom}}. \tag{10.6}$$

One may therefore be worried that the estimate of the conditional probability will be more noisy for small values of P_{pres} and P_{nom}, where small changes in one of these quantities will have a large impact on the value of the ratio. Figure 10.2 shows that CP estimates have indeed a higher variance for lower values of the probabilities. This indicates that the estimation of CP becomes very noisy when the probability of events is very small.

A natural consequence of this situation is that for estimates of CP with markets with low probabilities, it is hard to interpret short-term movements in CP values. Recent movements in the CP may be real movements in the fundamental CP value or just temporary movements due to the high sensibility of the CP estimate. One can naturally think of smoothing the estimation over time to have a more stable estimate of the CP. This would be equivalent to discounting the most recent evolution in CP to take into account the variability of the CP estimates in the short-term.

4 Biases

Low liquidity and higher variance for events with a low probability are problematic, but do not fundamentally jeopardise the value of CP estimates. The most important question is naturally about the calibration of these estimates: can we expect prediction markets' estimators of CP to be good estimators of the real CP?

The answer to this question naturally depends on the quality of the estimation of probabilities by prediction markets themselves. Numerous studies have found biases in prediction market prices, although usually of limited magnitude.

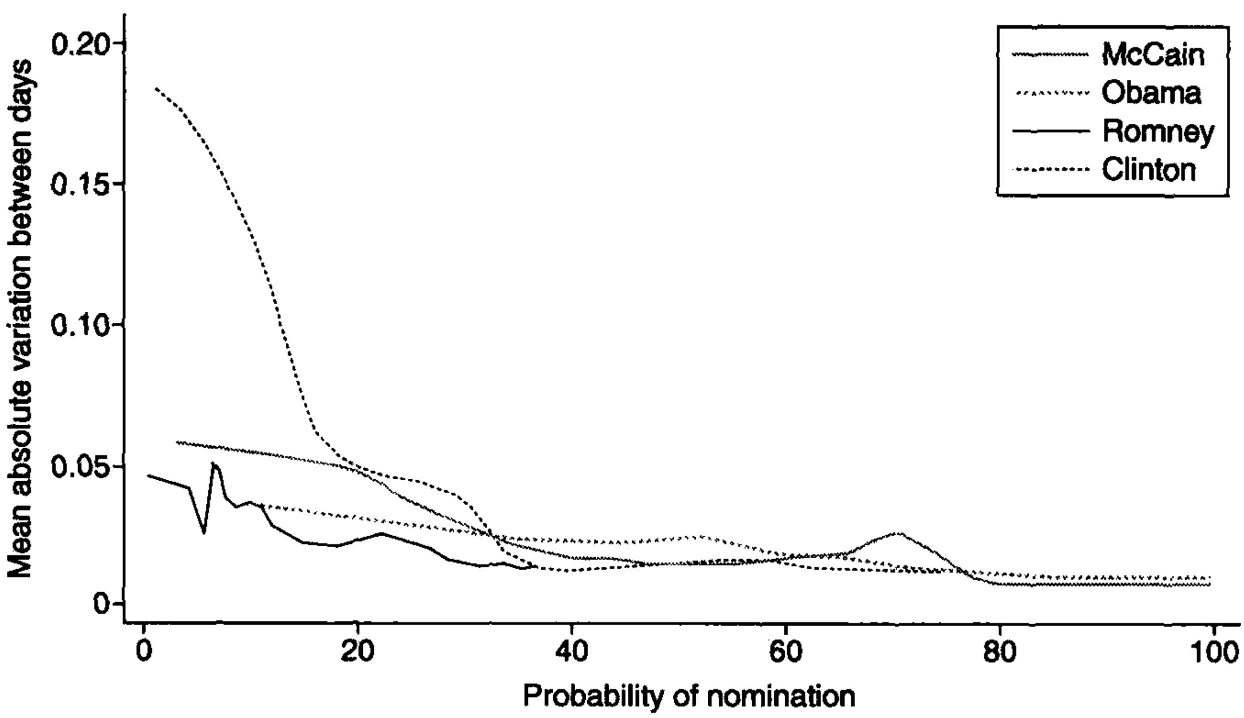


Figure 10.2 Variation of the CP as a function of the P_{nom} probability.

4.1 Short-term bias

Let's start from the situation where prediction market prices are relatively good estimates of the probability of an event. It is easy to see that a condition for the estimates of the CP to be unbiased is that the two market prices in the ratio integrate any new information at the same speed. If one market reacts quicker than the other, then the CP estimate will move in one direction for reasons unrelated to the underlying CP. Typically, a short-term bias will appear in the CP estimation if one market tends systematically to integrate new information more quickly.

Figure 10.3 shows how the CP estimates react to a short-term movement in the probability of being elected president. An increase in the probability of being elected president is associated with an increase in the CP. This could be considered as normal since it could actually be that the increase in CP raises the probability of being elected president. Figure 10.4 shows, however, that an increase in the probability of being nominated has a negative impact on the CP of being elected if nominated. This pattern is clearly unexpected. A candidate showing a better aptitude to win the primary race is a priori more likely to have the ability to win the presidential election. One would expect, if anything, a positive relationship between the probability of being nominated and the probability of winning the election if nominated. Figure 10.4 suggests the contrary.

A simple reason for this pattern is that the market on the primaries may integrate information about the primaries more quickly than the market about the

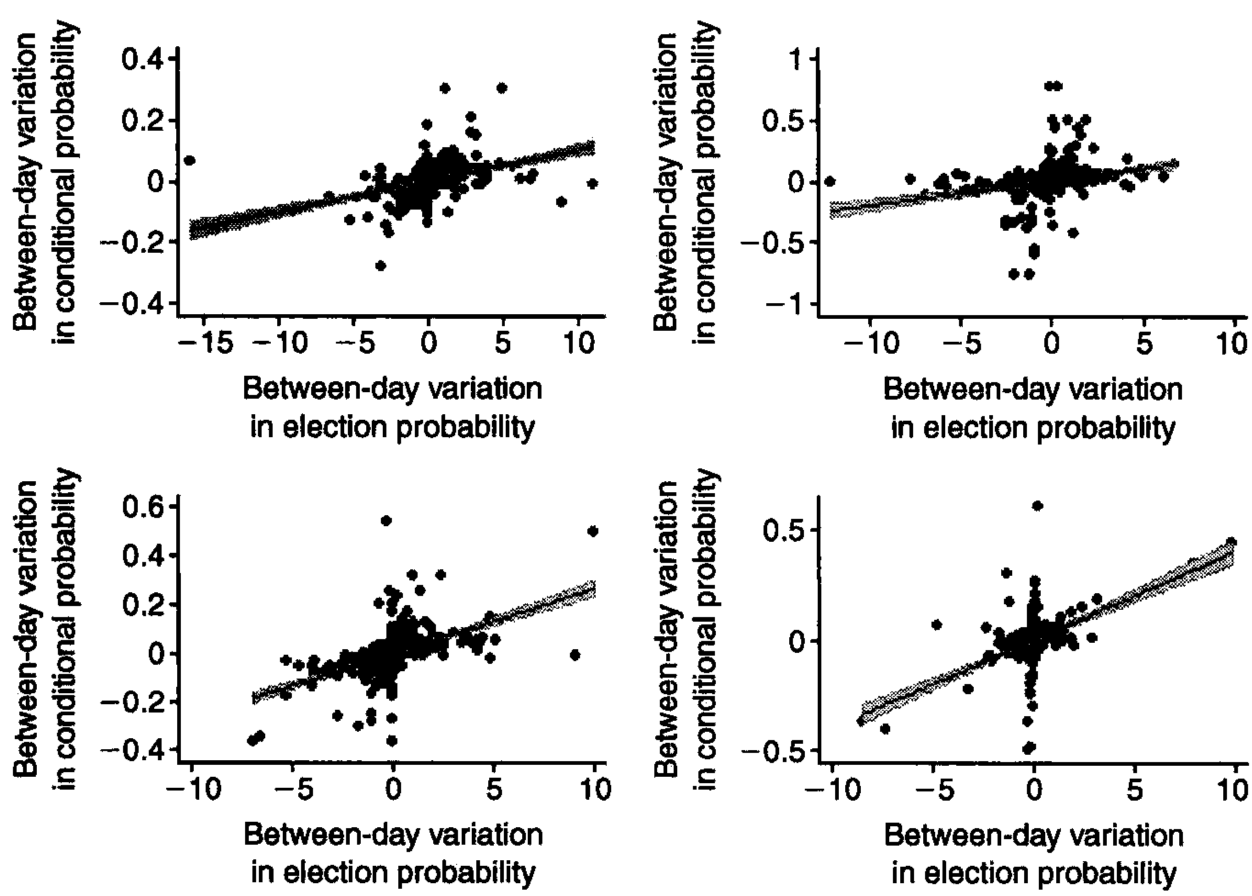


Figure 10.3 Effect of short term variation in P_{pres}.

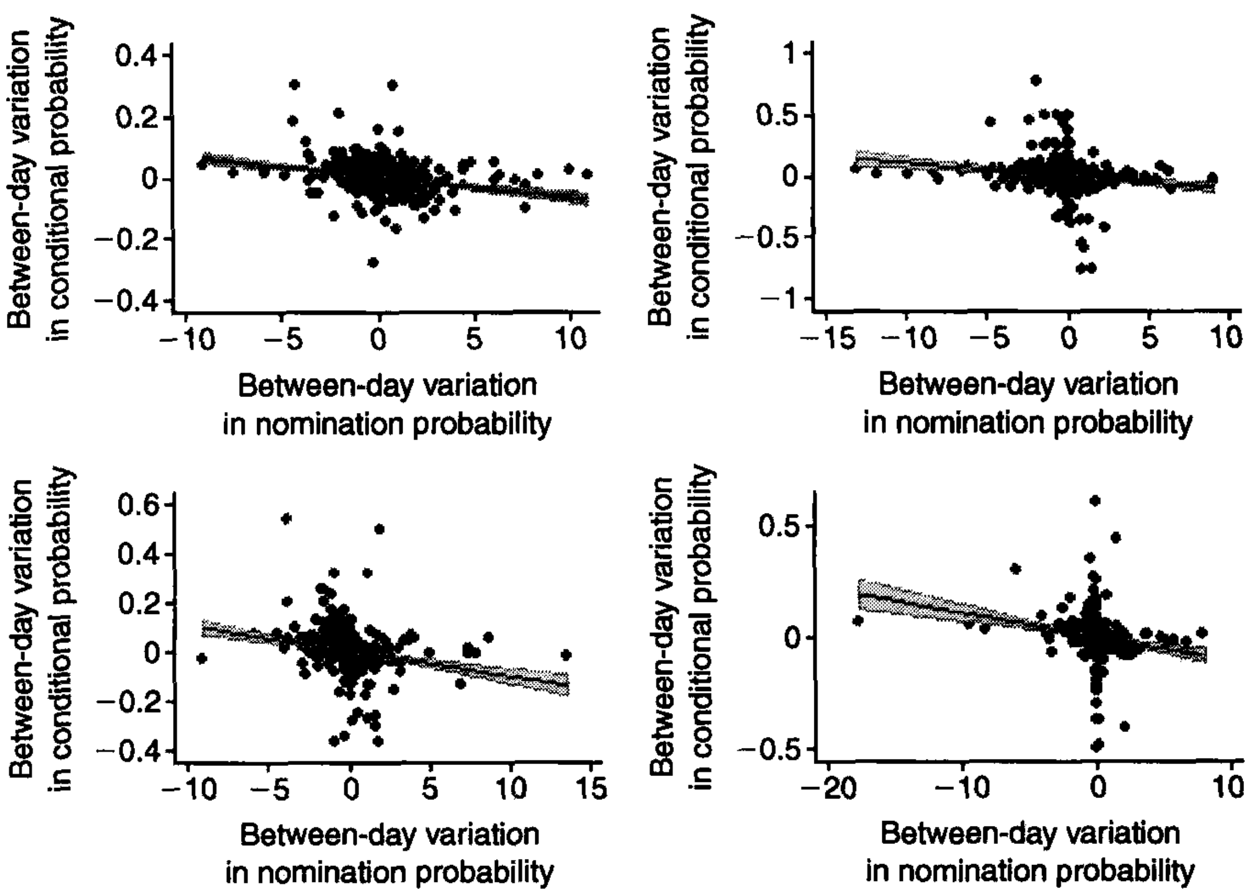


Figure 10.4 Effect of short term variation in P_{nom}.

presidential elections. Suppose a candidate wins a primary in one state, the market for his/her nomination may rapidly integrate this result into the price about his/her chances to be nominated. The market about the presidential election may take a bit more time to integrate this event into the probability of this candidate becoming president. This lag between the two markets will typically create a short-term bias such that a candidate who just won a primary will look less likely to win the presidency if nominated.

Another way to look at this phenomenon is to look at what happens to the CP estimate after a short-term shock in a price of the primary market. Figure 10.5 shows that after a negative or positive shock in the prices of the primary markets, the CP estimates react in a way predicted by the existence of a lag between the two markets. First, the CP estimate moves in the opposite direction from the primary market prices, and a day after it follows a readjustment in the other direction.

4.2 Longshot bias

Another natural concern is the existence of a longshot bias on prediction markets. A longshot bias means that prices overestimate small probabilities and underestimate high probabilities. If a longshot bias exists in each prediction market the CP estimate will be biased. Figure 10.6 shows the effect of a longshot bias for a CP of 50 per cent.

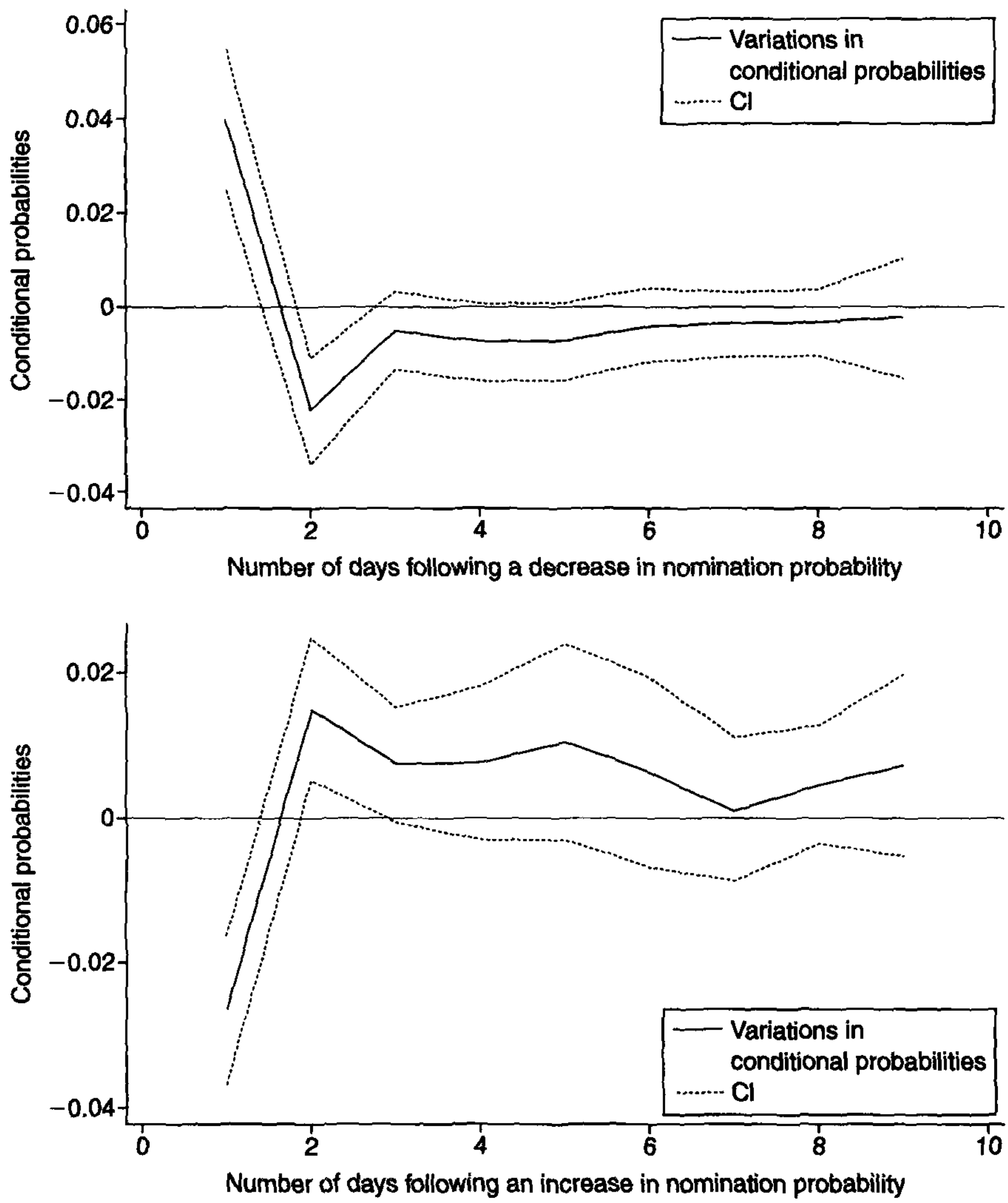


Figure 10.5 Bias and adjustment after a shock.

For an underlying probability p, the longshot bias is represented by a Lattimore *et al.* (1992) function:

$$price = \frac{p^{\gamma}}{p^{\gamma} + (1-p)^{\gamma}} \tag{10.7}$$

with a γ coefficient of 0.6 and 0.8. Figure 10.6 shows that a longshot bias on the prediction markets tends to systematically bias the estimation of the CP in two

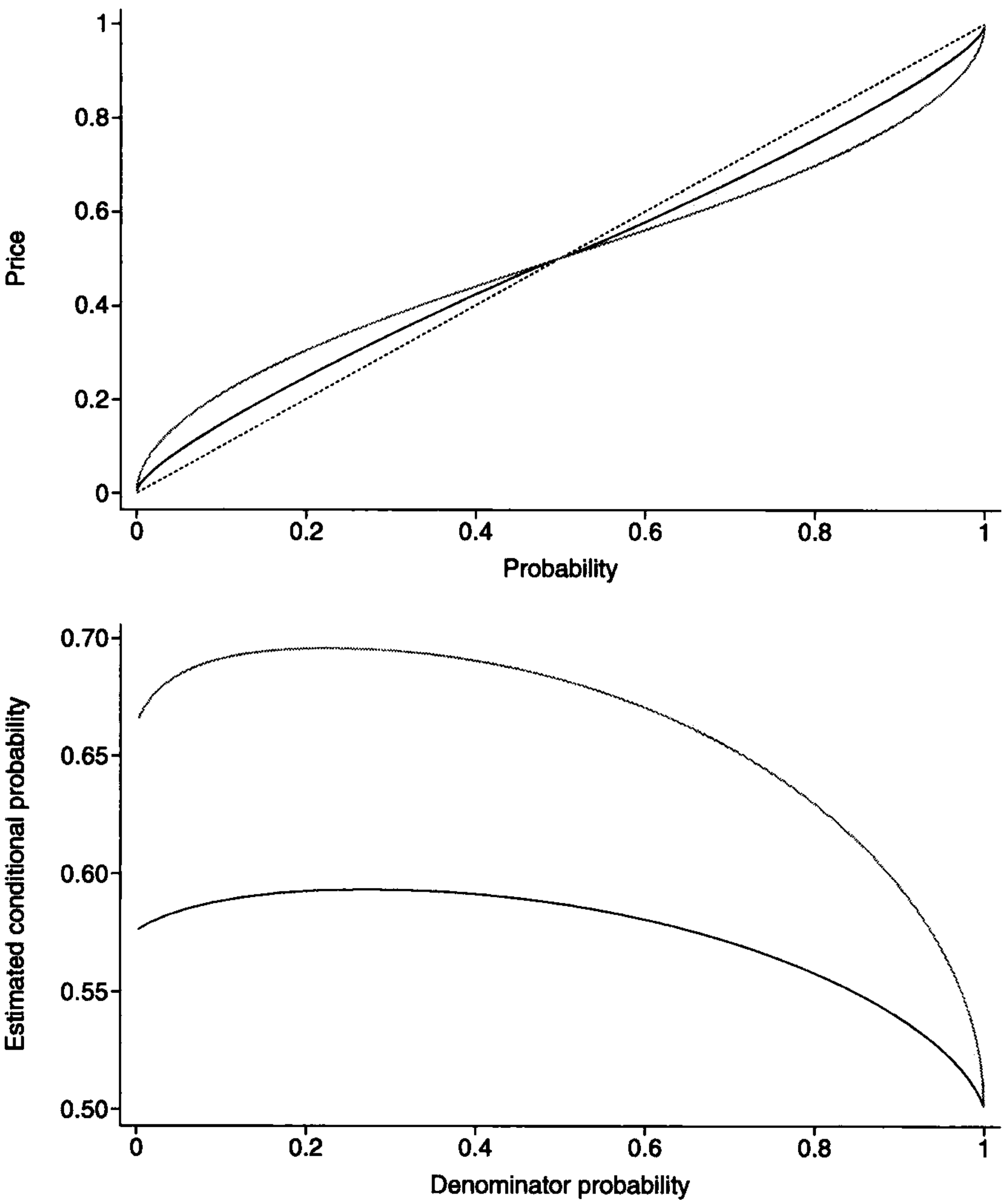


Figure 10.6 Longshot bias on a prediction market and estimation of a CP of 50 per cent. Estimation of a conditional probability of 50 per cent for $\gamma=0.6$ (up) and $\gamma=0.8$ (down).

ways. First, the estimation is higher than the value of the CP; second, the bias is higher for lower values of the denominator probability (here P_{nom}). This pattern is easy to explain. With the exception of the situations where the CP is close to 1, the probability of the joint event (here P_{pres}) will always be smaller than the probability of the denominator probability (here P_{nom}). As a consequence, due to the longshot bias, the price of the joint probability will be overestimated relative to the price of the denominator, which is higher. The estimation of the CP will therefore be overestimated. As Figure 10.6 shows, this bias can be very large, depending on the magnitude of the longshot bias.

The longshot bias creates another problem: the bias decreases with the level of the denominator probability. As a consequence, not only is the CP of each option biased, but the ordering of the CP for each option can also be biased. Suppose the decision-maker has to chose between two options, A and B. The decision-maker may not be too concerned that the estimate of the CP of success for each option is overestimated, but at least the relative ordering should be preserved. That is, if $CP(A) > CP(B)$, the decision-maker needs the estimates from the prediction markets to give $\widehat{CP}(A) > \widehat{CP}(B)$. The existence of a longshot bias actually breaks this relation as it is possible that while $CP(A) > CP(B)$, $\widehat{CP}(A) < \widehat{CP}(B)$ if the probability of A is higher than the probability of B. This is even more worrying if one thinks that it is somewhat likely that if A has a higher probability it may have also a higher probability of success.

It is possible to examine in our data set how, for a given candidate, the conditional probability evolves as a function of the probability of being nominated. Figure 10.7 shows the estimated CP as a function of the nomination probability using a fully flexible non-parametric estimation (local linear regression). The conditional probability is centred per candidate to give, in some sense, a non-parametric estimation of the effect of an increase in the nomination probability on the CP. As predicted by the existence of a longshot bias, the estimated conditional probability is actually higher for a low level of nomination probability.

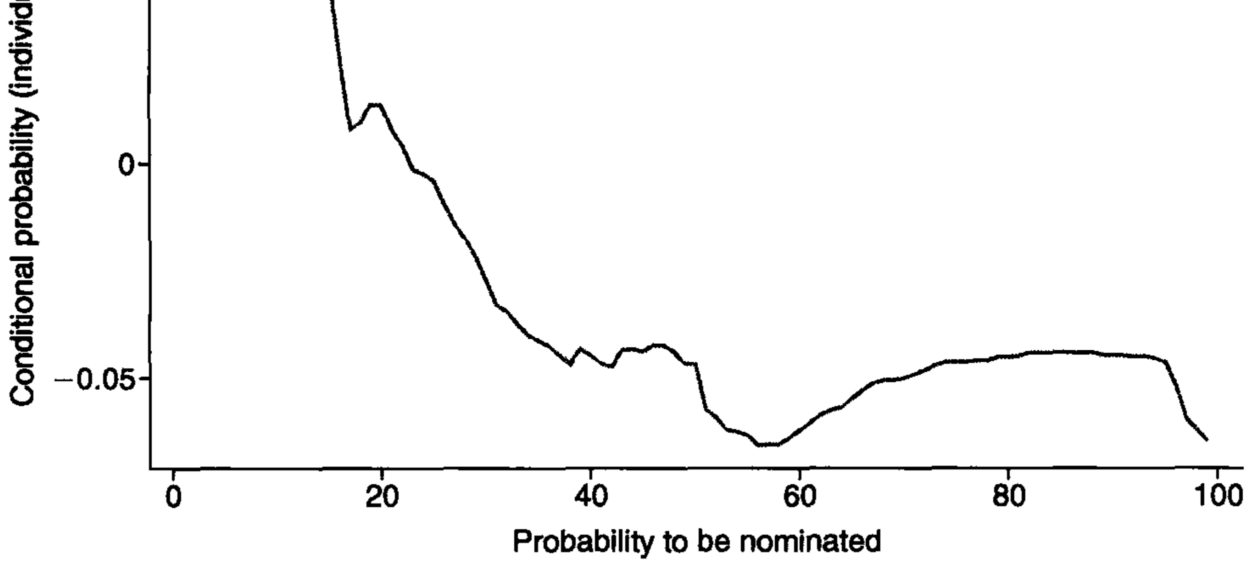


Figure 10.7 Estimated conditional probability relative to the nomination probability.

4.3 Manipulations

A final concern about the value of the CP estimate arises from the risk of manipulation. Numerous studies have addressed the question of manipulation on prediction markets without any conclusive results. Empirical and theoretical evidence is mixed on this issue. A field experiment by Camerer (1998) found a resilience of betting markets regarding manipulations. A historical study of manipulation attempts in prediction markets and a field experiment with manipulative bets of $500 by Rhode *et al.* (2006) suggests that manipulations have only a short-term effect on prices. Hanson *et al.* (2006) also found no effect on manipulation in their experiment. On the other hand, Hansen *et al.* (2004) report successful attempts at manipulating prices in the Iowa Electronic Markets in the field. From a theoretical point of view Hanson and Oprea (2000) proposed a model of market where manipulation ends up causing prediction market prices to be more accurate due to the liquidity they provide. This model, however, implies that, as in their 2006 experiments, traders know the existence of the manipulators and the direction and the incentive of the manipulator. These hypotheses may actually be inadequate to describe field prediction markets where traders may be unsure about the presence of a manipulation and about its direction. Recently, a case of successful manipulation on the price of McCain during the last stage of the US presidential campaign seem to have been observed on the highly liquid InTrade presidential market, with the price of McCain being raised by nearly ten points for more than one month (Rogin, 2008).

The concern about manipulation of CP estimates should be higher. First, because CP estimates are more important for decision-makers, there is a higher incentive to manipulate these markets in order to influence the decision-maker's choices. Second, due to the ratio properties of the CP estimator, a manipulator can benefit from a leverage effect to modify a given CP estimate:

$$\frac{\partial \frac{P_{pres}}{P_{nom}}}{\partial P_{nom}} = -\frac{P_{pres}}{P_{nom}^2} = -\frac{CP}{P_{nom}}, \tag{10.8}$$

If with an amount $x a manipulator can move the price by one percentage point, then for a CP of 50 per cent she could move the CP by two percentage points if the P_{nom} is 50 per cent, three percentage points if the P_{nom} is 33 per cent, and 5 percentage points if the P_{nom} is 10 per cent. This increased marginal effect of a given amount of manipulative money on the market is making it more likely for a manipulation to be successful. It also increases the probability of manipulation attempts by increasing their likelihood of success.

5 Meaning

Finally, one may wonder about the meaning of what we estimate.

First, in the case where there are some biases, one may be concerned that the changes observed are not changes in the underlying value but changes in the biases. For instance, there is evidence that longshot bias diminishes over time. One is therefore likely to observe a decreasing CP over time simply due to the decrease in the upper bias stemming from the longshot bias.

The problem of the meaning of the estimation is, however, deeper. Even in situations where we can think that the CP is correctly estimated by the market, one may wonder if we are measuring what we are really looking for from the decision-maker's point of view.

From the decision-maker's point of view, we are interested in $P(A|B)$ because we are interested in the *causal effect* of B on A. Even if the prediction markets can estimate $P(A|B)$, this is sometimes different from the causal effects. The problem arises from the fact that the prediction market can include in the CP estimate some information that we would like to exclude. For instance, the market may not estimate $P(A|B)$ but $P(A|B, P(B))$. To give an example, consider the US presidential elections. If Hillary Clinton had a very high probability of winning if nominated $P(W|N)$ when her probability of being nominated was very low, it could be because if she had won the nomination from her outsider position, she would had huge momentum. In this example the probability $P(N)$ itself influences the conditional probability of success.

Another possible problem is that the realisation of B may reveal some information having an influence on A. Suppose, for instance, that there are two different policies to tackle a problem and that a scientific committee has to give a report to suggest which policy is the best. The prediction market may anticipate the fact that if a policy is chosen eventually, it suggests that supportive evidence will have been found by the committee. In this case, calling F the set of information available at the time of decision between the two policies, the prediction market estimate may be $P(A|B, E(F|B))$. If B is eventually chosen it suggests that the set of information F is likely to be favourable to the success of B. This is naturally problematic as one would like to have an estimate of the efficiency of B estimated now, not with the inclusion of the expected positive evidence if B is chosen.

The two preceding examples may be linked. For instance, in the situation where Hillary Clinton has a very low probability of winning the primary election, a win for her in these elections would reveal that she was good at winning difficult electoral races and this would therefore indicate that she has a good probability of winning the presidency. In this precise situation the market would estimate $P(A|B, E(F|B, P(B)))$, i.e. the probability of winning if nominated takes into account both the expected information on the candidate for presidency if nominated and the knowledge that she had initially only a low probability of success.

Clearly this is not really what a decision-maker (here a Democrat voter) would like to know; instead, a decision-maker would like to compare for a given state of the world F, $P(A|B, F)$ versus $P(A|B^C, F)$. That is, what is the effect of changing B for not B, based on the probability of A happening. Here, for instance, the Democrat voter would be interested in the probability of changing Clinton for Obama based on the probability of winning the presidential election.

Prediction markets may actually incorporate too much information in the estimation for it to be safely considered as representing the causal effect of *B* on *A*.

6 Conclusion

Prediction markets can provide a way to estimate conditional probabilities. Given the importance of conditional probabilities for decision-making processes, this is surely something to study in depth to see how well prediction markets estimate CP and how these estimates can be improved. I have shown in this chapter that prediction markets face a number of difficulties when trying to estimate conditional probabilities as a ratio of two separate prices. These estimates present a high variance for small levels of probabilities, a short-term bias due to a lag in the adjustment of each market, and an in-built upward bias due to the effect of the longshot bias. In addition, the meaning itself of prediction market estimates may be problematic as the estimate may actually incorporate too much information to represent exactly the causal effect that the decision-maker is looking for in order to make her decision.

Ways around these caveats should be looked for, but there is no easy solution. We could mention in particular here the idea of 'conditional prediction markets'. A conditional prediction market is a market whose contract is valid if an event occurs. We could have, for instance, a contract on '*A* if *B*'. This contract will be valid if *B* occurs (and be worth 0 or 100, depending on the occurrence of *A*), but the contract will be void if *B* does not occur (in this case the traders would be reimbursed for their investment). Such markets have never been seriously implemented in practice, so it is hard to predict how well they would predict probabilities. However, there is at least one simple problem for these markets, namely that the degree of commitment in the contract decreases with the probability of *B*. This could make the market more prone to manipulation. If, for instance, a candidate has only a 5 per cent chance of winning the primaries, a manipulator betting on the conditional market for this candidate to win the presidential elections if nominated would only have exposure to 5 per cent to see the contract realised. There would therefore be less risk associated with manipulation activities to raise upward the conditional probability price. More fundamentally, one could expect that a larger proportion of trades would not be made 'rationally' due to this lower level of commitment in trading activities.

References

Berg, J. and T. Rietz (2003) 'Prediction Markets as Decision Support Systems', *Information Systems Frontiers*, 5(1), 79–93.

Camerer, C. (1998) 'Can Asset Markets be Manipulated? A Field Experiment with Racetrack Betting', *Journal of Political Economy*, 106(3), 457.

Hansen, J., C. Schmidt and M. Strobel (2004) 'Manipulation in Political Stock Markets: Preconditions and Evidence', *Applied Economics Letters*, 11(7), 459–463.

Hanson, R. (1999) 'Decision Markets', *IEEE Intelligent Systems*, 14(3), 16–19.

Hanson, R. (2007) 'Shall We Vote on Values, But Bet on Beliefs?', George Mason working paper.

Hanson, R. and R. Oprea (2009) 'Manipulators Increase Information Market Accuracy', *Economica*, 76, 304–314.

Hanson, R., R. Oprea and D. Porter (2006) 'Information Aggregation and Manipulation in an Experimental Market', *Journal of Economic Behavior and Organization*, 60(4), 449–459.

Henderson, M., J. Wolfers and E. Zitzewitz (2008) 'Predicting Crime', working paper.

Lattimore, P., J. Baker and A. Witte (1992) 'The Influence of Probability on Risky Choice: A Parametric Examination', NBER working paper.

Mankiw, G. (2008) 'McCain's Likely Tax Policy'. Online: http://gregmankiw.blogspot.com/2008/09/mccains-likely-tax-policy.html.

Page, L. and B. Clemen (2008) 'Are prediction markets well calibrated prediction tools?', working paper.

Rhode, P., K. Strumpf, U. Hill and N. Hill (2006) 'Manipulating Political Stock Markets: A Field Experiment and a Century of Observational Data'. Online: www.unc.edu/~cigar/papers/ManipNBER.pdf.

Rogin, J. (2008) 'Trader Drove Up Price of McCain Stock in Online Market', *Cqpolitics*. Online: www.cqpolitics.com/wmspage.cfm?docID=news-000002976265.

Wolfers, J. and E. Zitzewitz (2004) 'Experimental Political Betting Markets and the 2004 Election', *The Economist's Voice*, 1(2), 1–8.

11 Prediction markets

A study on the Taiwan experience

Shu-Heng Chen, Chen-Yuan Tung, Chung-Ching Tai, Bin-Tzong Chie, Tzu-Chuan Chou and Shu G. Wang

1 Motivation and introduction

1.1 History of prediction markets in Taiwan

Among all Asian countries, Taiwan is probably the first one to have her own prediction markets. From the early 2000s to the mid-2000s, three prediction markets were established one after the other in Taiwan. This happened a few years before the first political prediction market in Japan, General Election Hatena, which was established in 2005 by Hatena Co.,[1] and the first prediction market in New Zealand, ipredict, which was established in 2008 by Victoria University of Wellington and the New Zealand Institute for the Study of Competition and Regulation (ISCR).[2] In this section, we shall give a brief introduction to the Taiwan prediction markets, their history, operation, research and publications.

The Taiwan Political Exchange (TAIPEX) was established in 2003 at the Institute of Physics, Academia Sinica in Taipei. Since then it has functioned several times to predict many important political events, including both the US and Taiwan presidential elections in 2004 (Wang *et al.*, 2004, 2006, 2009).[3]

The AI-ECON Futures Exchange (AI-ECON FX) was established in 2006 at the AIECON Research Center, National Chengchi University in Taipei.[4] It was established in an attempt to integrate *agent-based computational economics* and *experimental economics* as an initial step to further overarch *computational social sciences* with *experimental social sciences* (Chen and Tai, 2006; Barr *et al.*, 2008). This project was sponsored by the National Science Council for three years from the middle of 2006 to the middle of 2009. AI-ECON FX has been applied to predict the opening day of the high speed railway in Taiwan. In November 2006 both TAIPEX and AI-ECON FX were run in parallel to predict the Taipei and Kaohsiung City mayoral elections (Chen and Wu, 2009).[5]

In addition to predicting future events, AI-ECON FX functions both as an *experimental market*[6] and a *computational platform*. In terms of the former, it helps to observe and study the behavior of market participants. Two empirical issues have already been addressed by using TAIPEX and AI-ECON FX. The first issue concerns the *evolving network topology* observed in the prediction markets (Wang *et al.*, 2008);[7] the second issue is the *bidding strategies* of market participants (Chen and Hsieh, 2010). These two issues are fundamental because they can shed light on what is known as the *Hayek hypothesis* (Smith, 1982) or the mythical *aggregation mechanism*, which is frequently claimed to be the theoretical foundation of prediction markets. The latter helps to construct and simulate the *agent-based prediction markets*. Specifically, we ask whether we can design artificial agents to replicate human behavior observed in prediction markets. The cases in point are Tseng *et al.* (2008, 2009), who use artificial agents to replicate some features of the network topologies generated by human traders in TAIPEX and AI-ECON FX.

However, in terms of scale, scope and activeness, neither TAIPEX nor AI-ECON FX are comparable to the Exchange of Future Events (XFuture), which is jointly run by the Center for Prediction Markets, National Chengchi University and SWARCHY.[8] XFuture was initially set up in 2006. It has since boomed in popularity, particularly following the setting up of the Center for Prediction Markets in mid-2006.[9] Four years on (up to July 2010), XFuture has issued 14,938 futures distributed over 2,115 events with 1.98 million submissions, 98,000 matches and a total trading volume of 240 million. Participants are no longer restricted to local people. Registered traders come from 4,842 cities in over 121 countries, which covers all cities in Taiwan, 858 cities in China and 1,912 cities in the United States.

While XFuture was originally designed for election prediction, it has gradually developed into a comprehensive market for various kinds of uncertain events, including politics, economics and finance, interior affairs, cross-Strait (China–Taiwan) affairs, international affairs, sports and entertainment. In addition to a number of nation-wide elections in Taiwan and the key local and international financial indexes, most traded predictions also include some important policies, such as the opening-up time of Taiwan to tourists from the Chinese mainland.

XFuture is distinguished by its market design. Unlike most prediction markets, XFuture runs as a futures market in the sense that market participants need only pay a futures margin while trading. The calculation of the futures margin is also unique, and is different from the ones normally applied in futures markets. The basic principle of calculating a futures margin in XFuture is based on the worst-case assumption. In the case of the winner-takes-all markets, the margin can be figured out quite easily because the termination price is either 100 or 0. Hence, when a trader shorts for ten units at a price of 30, the worst case that could happen to him is a termination price of 100. In this case, he has to cover the short with a total of 700, i.e., $(100-30)\times 10$. The margin is, therefore, 700.

1.2 Successes and failures in XFuture

Very few studies of XFuture are, however, available. Based on the two recent nation-wide elections, Tung *et al.* (2009) show that XFuture exhibits superior performance to the polling institutions over a number of criteria of forecasting accuracy. This chapter is a continuation of further exploration of the XFuture database. We consider a collection of 575 futures, which belongs to 172 political events (see Table 11.1). The common feature they share is that they are all issued in the "winner-takes-all" market.[10] Among these 172 events with 575 futures, 151 have made correct predictions. The hit rate is, therefore, as high as 87.79 percent.[11]

Our experience of these 172 events with 575 futures naturally compels us to wonder why some events are predicted well but others are not. The research questions on the successes or failures of prediction markets have been quite well discussed in the literature. In a more general context, Wolfers and Zitzewitz (2004), in their introductory review of the literature on prediction markets, spell out the two essential ingredients pertaining to *market designs*: roughly speaking, the effectiveness of the *aggregation algorithm* and the quality of *market participants*. As implied, an ideal design is a well-articulated contract with regard to a reasonably interesting or entertaining issue circulated in a continuous double-auction market in which conflicting information and heterogeneous beliefs are widely held by market participants. In a separate review of the work on Iowa political futures markets, Berg *et al.* (2008) reiterate these two ingredients:

> For the markets to work in theory, two features must be present. First, there must be enough traders so that the aggregate of their knowledge can forecast correctly the outcome of the election. Second, the market mechanism must facilitate aggregation of their disparate information so that the prevailing market price becomes a sufficient statistic for the traders' collective information.
>
> (Ibid.: 748)

They further identify three factors which explain most of the variance in the accuracy of the Iowa Election Market (IEM) predictions (ibid.: 746):

1 Presidential election markets perform better than (typically lower profile) congressional, state and local election markets.
2 Markets with more volume near the election perform better than those with less.
3 Markets with fewer contracts (i.e., fewer candidates or parties) predict better than those with more.

These three points are directly applicable to our dataset with 172 political events. It is, therefore, interesting to know whether we can use these factors to separate our 172 events into two groups, one which successfully predicted the outcome, and the other comprising those which failed to do so. The factors together suggest the relevance of the following respective variables: (1) the popularity of the political events; (2) trading volume near the election; and (3) the number of options. While it has not been mentioned explicitly, the number of traders is obviously relevant, since it is the basis on which information is aggregated. The contribution of these four variables,

Table 11.1 Events and contracts used in the analysis

Year	*Nature*	*Number of defined events*	*Number of contracts*
2006	Municipality The biggest party in the municipal congress (local)mayoral election (local)	2	10
		2	6
2007	Lawsuits (local)]	1	14
	Presidential nominations (national)	4	16
	Legislative election (national)	2	2
2008	Legislative election (national)	76	292
	Presidential election (national)	2	4
	Referendum (national)	2	2
	Presidential election, winners in each area (local)	25	75
	U.S. presidential nomination, Democratic Party (foreign)	1	3
	U.S. presidential election (foreign)	3	5
	New party leader (national)	1	4
2009	Legislative by-election (local)	9	50
	County magistrate election (local)	35	67
2010	Legislative by-election (local)	7	25
Total		172	575

with some of their extensions and variations to be detailed in Section 2, will then be examined using different quantitative models.[12]

The rest of the chapter is organized as follows. Section 2 provides a general description of the data and the variables that are used to distinguish successes from failures. Section 3 introduces the quantitative models used to distinguish success from failures. Section 4 presents the empirical results. Concluding remarks are given in Section 5.

2 Data description

2.1 General description

Before proceeding with further quantitative analysis, we shall give a general description of our data and the chosen input variables. Our data set is composed of 172 political events of which the time of occurrence lies between 2006 and 2010. The characteristics of these 172 political events are all given in the second column of Table 11.1 and are distinguished by the year of occurrence (expiration), shown in the first column. For example, there are two kinds of political events underlying XFuture in 2006: mayoral elections and the biggest party in the municipal congress. As also indicated in the second column (inside the parentheses), all involved events are local events. Moving further to the right of the table, we provide the number of events (the third column) and the total number of futures contracts (the fourth column). So, to continue the example for 2006, there are a total of ten futures contracts for two mayoral elections, and six futures contracts for two dominant-party events. A further breakdown may help us see the relationship between columns three and four. As for the two mayoral elections in 2006, one was held in Taipei and the other in Kaohsiung. The former had six candidates, and the latter had four. Hence, they are summed to ten.

A quick glance at Table 11.1 reveals that most political events mentioned are election-oriented, and almost all of them refer to elections in Taiwan, be they national or local. However, there are also a few foreign political events involved, and they are either directly or indirectly related to the US presidential elections. For example, one indirect event inquires about the impact of Barack Obama (or John McCain) being elected in the 2008 US presidential election on the Dow Jones Index.

Earlier we mentioned that the 575 future contracts of the 172 events are all issued and traded in winner-takes-all markets. In the winner-takes-all design, each event is associated with one or more than one option (futures). We therefore distinguish those events that have only one option from those that have two or more, i.e., the *single-option events* vs. the *multiple-option events*. In the case of the single-option events, the value of the futures contract is interpreted as the probability of a specific event occurring, say, *A*. The alternative, *B*, is not specified and not traded as an option. In this case, we say that the market correctly predicted the future if the closing price of the futures contract associated with *A* is greater than 50.[13] An example is Event 56, "the first-ranking at-large seat of the KMT for the national legislative election in 2007 is Jin-Pyng Wang." Since this event does not specify who will be the alternative if not Jin-Pyng Wang, it is considered to be a single-option event. This futures closed at a price of 99.80, a price far greater than 50, which implies that Jin-Pyng Wang would almost certainly by the first-ranking at-large seat of the KMT in 2007. Then, indeed, the event came true. Hence, this prediction is correct.

Most events, however, have multiple options. The typical example is an election involving a number of candidates, in which we have a futures contract for each candidate. In this case, the future (candidate) with the highest price is considered to be the market prediction. Take the 2008 presidential election in Taipei City, Taiwan (Event 218) as an example. There are three options under this event; they correspond to the two major candidates, Ying-Jeou Ma and Frank Hsieh, and others, respectively. The three options are stated as follows: "Ying-Jeou Ma's vote share is the highest among all candidates in Taipei," "Frank Hsieh's vote share is the highest among all candidates in Taipei," and "Neither Ma's nor Hsieh's vote share is the highest among all candidates in Taipei." The closing price of these three futures are 99.44, 0.50, and 0.10, respectively. Hence, based on the highest price, the market predicts that the candidate with the highest vote share in Taipei City is Ying-Jeou Ma. This prediction turns out to be correct because Ma's vote share in Taipei city is 63.03 percent and Hsieh's vote share is only 36.97 percent.

2.2 Attributes to distinguish between successes and failures

We can classify these 172 events into the events that were correctly predicted (success events) and the events that were not (failure events). We then try to see whether we can distinguish between the two classes by using some variables suggested by the literature. Based on the earlier discussion (Section 1), we have included a list of the following variables.

Number of contracts under the same event. Due to Wolfers and Zitzewitz (2004), this variable is included. Figure 11.1(a) gives the histogram for this variable. Table 11.2 also gives the basic statistics for this variable. The median size has three futures contracts (three options), but the maximum one has 14 contracts.

Number of active traders. Not all registered traders are active. Those who were idle obviously did not contribute to information aggregation. Therefore, here we only count those who made submissions, while not necessarily being matched successfully. A histogram of the number of active traders is given in Figure 11.1(b). A wide range of this value exists from the thin markets to the thick markets (Table 11.2). The thinnest market has only five active traders, whereas the thickest market has more than 7,000. The median size of the active traders is 63.5.

Closing trading volume. Trading volume not only measures how active the market is, but can also be an indicator of the degree of the heterogeneity of the market. Basically, transactions occur because agents hold heterogeneous beliefs

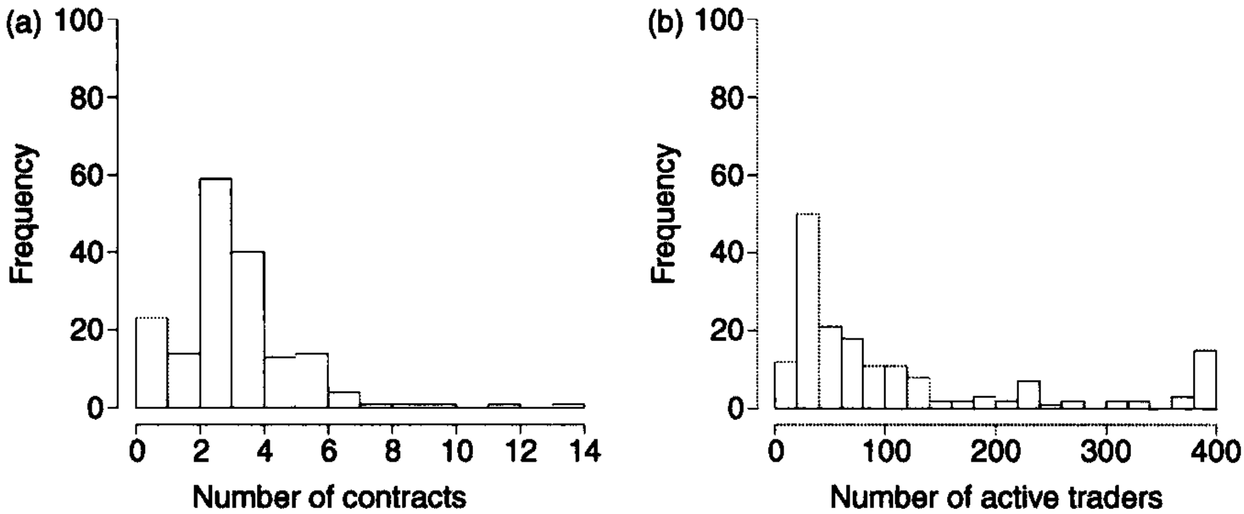


Figure 11.1 Number of futures contracts (a) and number of active traders (b).

Notes
The histogram of the number of active traders (b) is only drawn up to 400 so as to maintain its visibility. Markets with more than 400 active traders are then included in the last block of the histogram. That is why we see a spike appearing in the tail.

or different information. To make the prediction market work, it is desirable to have an aggregation over a set of heterogeneous agents so that different sources of information can be effectively pooled together. Given the possibility that the state of the market can be constantly changing due to the arrival and dispersion of new information, we therefore decide to take the trading volume of the last day as a proxy for the measure of heterogeneity. Figure 11.2(a) gives the histogram for the closing trading volume. The difference between the thin markets and the thick markets can also be seen in this figure. The closing trading volume varies from one unit as the minimum to 1.4 million units as the maximum, and has a median of 587 (Table 11.2).

Open interest. As we have discussed above, XFuture runs as a futures market; therefore, an additional variable that is not generally shared by other prediction markets is *open interest.* Open interest is the measurement of those traders in the

Table 11.2 Descriptive statistics of the attributes

	Number of contracts	*Number of active traders*	*Volume*	*Open interest*	*Number of traders with open interest*
Mean	3.59	233.22	24,880.74	119,523.48	223.22
Median	3	63.5	587	4,296	68
Min	1	5	1	14	4
First Quartile	3	32	59	726	32
Second Quartile	4	135.25	6,812.5	29,311.25	127.75
Max	14	7,396	1,403,635	6,610,828	5,360

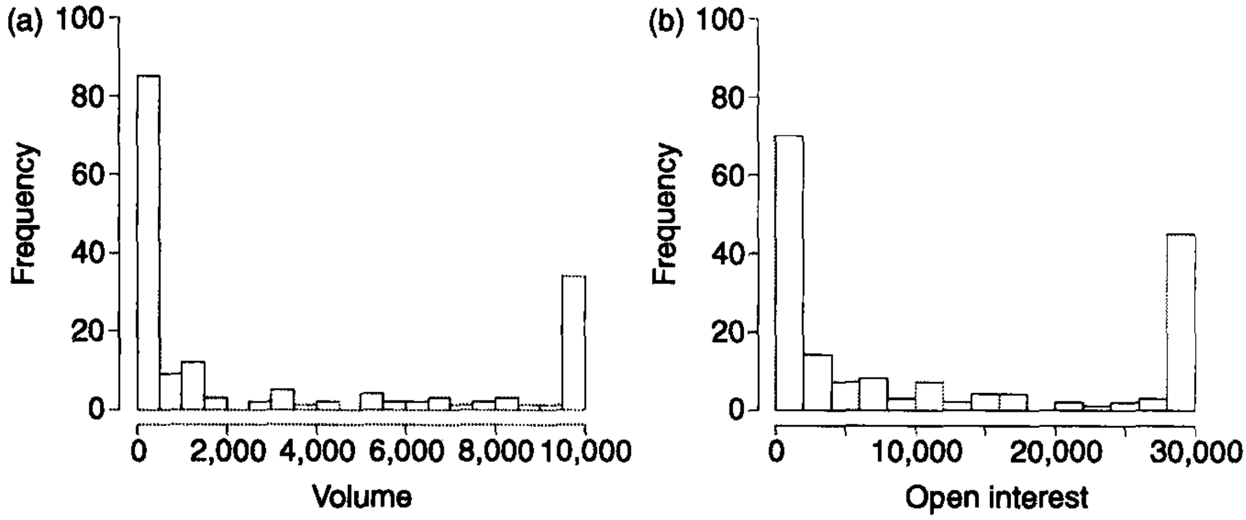


Figure 11.2 Closing trading volume (a) and open interest (b).

Note
The tails of both histograms spike for the same reason as provided for Figure 11.1.

futures market with outstanding trades. Open interest is the net value of all open positions, be they short or long, in one market. If open interest increases suddenly, it is likely that new information has been revealed. A market with a low trading volume but a large open interest indicates that there are many participants who will enter the market only when the price is right. Therefore, open interest can be used as a measure of the rightness of the market price. Figure 11.2(b) gives the histogram of open interest. Huge divergence also exists in this variable, ranging from 14 units to 6.6 million units, with a median of 4,000 (Table 11.2).

Number of traders with open interest. The contribution of the intra-marginal traders to price discovery has been noted in the literature. The existence of a large number of intra-marginal traders, who are able to either bid higher or ask lower than the equilibrium price, enhance the matches of buyers and sellers and hence the price discovery. However, to distinguish the intra-marginal traders from extra-marginal traders is not easy. Nevertheless, it is our conjecture that open interest may be relevant to differentiate between the two types of traders. Therefore, we consider the number of traders with open interest. Figure 11.3(a) gives the number of traders with open interest. Since not all traders will leave their position open on the final day of the market, this variable has a narrower distribution than the number of active traders, with a range from 4 to 5,360 and a median of 68 (Table 11.2).

Levels of political events. The levels of political events are included because they may be associated with different kinds of information dispersion and hence different operations of aggregations. In the national events, the acquired information among traders is expected to be more heterogeneous; in addition, the presence of social and geographical blocks can be more evident. Therefore, the aggregation with this greater heterogeneity and segregation can be more difficult. On the other hand, national events can arouse more interest among the public and hence

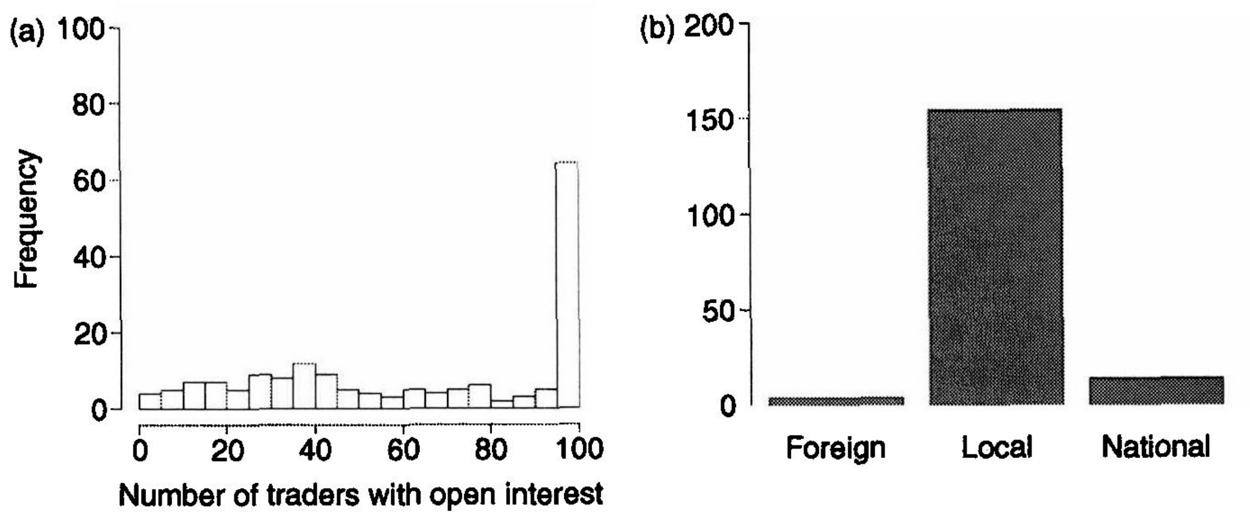


Figure 11.3 Number of traders with open interest.

Note
The tail of the histogram spikes for the same reason as provided for Figure 11.1.

facilitate the formation of a more active market (a large number of traders with intensive trading), which may make aggregation more effective. The final net effect depends on the interplay of these two countervailing forces. As we show in Figure 11.3(b), there are 14 national events and 154 local events. Out of the 14 predictions of national events, only one is missed (an accuracy rate of 92.85 percent); by contrast, out of the 154 predictions of local events 20 are missed (an accuracy rate of 87.01 percent). This very rudimentary analysis did indicate that the predictions of local events are less accurate than the predictions of national events.

The six variables above are, of course, not exhaustive. There are many other variables that we consider equally interesting. We will highlight some of them in the concluding section.

3 Quantitative models

The possible influence of the six variables discussed in the previous section will be analyzed using some familiar statistical models, including logistic regression, support vector machines and decision trees. The logistic regression serves as a baseline, and it gives us a quick grasp of the possible influence of each attribute on prediction accuracy (Section 3.1). Support vector machines then serve as non-linear extensions of the logistic regression (Section 3.2). However, it would be harder to derive rules from support vector machines; therefore, decision trees are used as companions to support vector machines (Section 3.3).

3.1 Logistic regression

Logistic regression is a standard tool for prediction probability. We use it as a starting point (a benchmark) for our analysis. The logistic regression can be written as

$$H = \frac{1}{1+e^{-Z}} \tag{11.1}$$

and

$$Z = \beta_0 + \sum_{i=1}^{5} \beta_i X_i + \sum_{i=1}^{2} \gamma_j D_j \,, \tag{11.2}$$

where $H \in [0, 1]$ and $X1, X2, \ldots, X5$ refer to the first five variables introduced in Section 2.2, following the same sequence, i.e., number of contracts, number of active traders, closing trading volume, open interest and number of traders with open interest. The last variable, level of political event, is introduced to the model by two dummy variables. $D1$ is the dummy variable for local events, and $D2$ is the dummy for national events.

3.2 Support vector machines

A support vector machine is a kind of neural network, which has been frequently applied to classification. It non-linearly maps an n-dimensional input space into a high dimensional feature space.

$$\varphi : V^n \to V^m, \tag{11.3}$$

where V^n is an n-dimensional input vector space, and V^m is an m-dimensional feature vector space. Given a series of l historical observations (in our case $l=172$):

$$(y_1, x_1), \ldots (y_1, x_1), \tag{11.4}$$

where $y_i \in V^1$ and $x_i \in V^n$; in our case $n=5$ and $x_i = (x_{1,i}, x_{2,i}, x_{3,i}, x_{4,i}, x_{5,i})$. We approximate and estimate the functional relation between y_i and x_i by:

$$y_i = f(x) = < w, \varphi(x) > +b = \sum_{j=1}^{m} w_j \varphi(x)_j + b, \tag{11.5}$$

where $<.,.>$ denotes the inner product. The vector w and the constant b are to be determined by following the *structural risk minimization principle*, borrowed from statistical learning theory (Vapnik, 1998). In our case, where the y_i are categorical, such as $y_i \in \{-1, 1\}$ (failure or success), the minimization process also determines a subset of $\{x_i\}_{i=1}^{l}$, referred to as *support vectors*, and the SVM when constructed has the following form:

$$f(x) = \sum_{s} y_i \alpha_i^* < \varphi(x_s), \varphi(x) > +b^* \,, \tag{11.6}$$

where α_i^* and b^* are the coefficients satisfying the structural risk minimization principle, and s is the set of all support vectors. The category assigned to the observation x, 1 or −1, will then be determined by the sign of $f(x)$:

$$y = \begin{cases} 1, & if \quad f(x) > 0 \\ -1, & if \quad f(x) < 0 \end{cases}. \tag{11.7}$$

Equations (11.6) and (11.7) are the SVM for the classification problem. A central concept of the SVM is that one does not need to consider the feature space in explicit form; instead, based on the Hilbert–Schmidt theory, one can use the *kernel function*, $K(x_s, x)$, where

$$K(x_s, x) = < \varphi(x_s), \varphi(x) > . \tag{11.8}$$

Therefore, the SVM is also called the *kernel machine*. Equation (11.6) can then be rewritten as:

$$f(x) = \sum_s y_i \alpha_i^* K(x_s, x) + b^* . \tag{11.9}$$

3.3 Decision trees

The decision tree has become a canonical tool in machine learning. It is a *classification* procedure with a tree-like graph structure. The data $S(=\{y_i, x_i\}_{i=1}^{172})$ presented to the decision tree is of a common type, namely, *six attributes* and *one decision class* (success or failure). Each attribute $A_j (j = 1, \ldots, 6)$ partitions the 172 inputs into s_j distinct classes based on the attribute value:

$$A_j : S \rightarrow \left(a_{j,1}, a_{j,2}, \ldots a_{j,s_j}\right). \tag{11.10}$$

When the input x_i is presented to the tree, at each node of the tree, a decision is made based on a test on a value of an *attribute*, a_j. According to the results of the test, the interpretation of the tree proceeds to one of the sub-trees of the node. The path will continue leading x_i to the next test until it goes through all of them, and hence reaches a leaf of the tree. It is expected that all paths of the decision tree will inform us better on how different decisions are made.

A decision tree is constructed based on a *top-down greedy algorithm*, known as the ID3 in machine learning (Quinlan, 1986). The idea is fairly straightforward. One first finds the attribute that *best* classifies the training data, and then uses this attribute as the *root* of the decision tree. Then the process is repeated for each sub-tree. The main issue involved in this greedy algorithm is the criterion regarding the choice of the best classifying attribute. A common solution to this problem is to select the attribute with the *highest information gain*, $G(S, A)$, which is defined as the expected reduction in the *entropy* of the data set S, caused by knowing the value of the attribute A.

An important issue pertaining to growing decision trees is *when to stop*. Would it be desirable to grow the decision tree until it perfectly matches the data? To avoid *overfitting*, the answer is generally "no." However, in practice, the greedy algorithm will grow the full tree first, then be pruned later. There are two different types of pruning. The first one is to prune the tree directly, known as *reduced error pruning* (Quinlan, 1987); the second is to first convert the tree into rules, and then to prune (generalize) each rule independently by removing preconditions that increase classification accuracy. This can be done by the famous algorithm C4.5 (Quinlan, 1993). In addition to pruning, one can also use a complexity measure, such as the *minimum description length* (MDL) to halt tree growth when the MDL is found (Grunwald, 2007).

4 Empirical results

4.1 Logistic regression

The results of the logistic regression are given in Table 11.3. None of the explanatory variables are statistically significant at the 0.05 significance level. However, five variables are somewhat significant at the 0.1 level. These are: the number of contracts, the number of active traders, closing trading volume, open interest and the number of traders with open interest. Even though we have seen that the prediction associated with national events has a higher degree of accuracy, the two dummies are not statistically significant in this multiple regression model. Using the regression equation obtained above to fit y_i (correctness of prediction markets), one obtains an accuracy rate of 88.95 percent, just marginally better than the guess "all predictions are correct." However, the logistic model has a poor capability to specify the failures (Table 11.4). Only three out of a total of 21 failures have been correctly singled out, which is very low specificity.

If we look further at the signs of the coefficients in Table 11.3, not all of them are expected. The sign of the number of contracts is negative, indicating that the higher the number of contracts, the less likely the correctness of the prediction. This finding is consistent with the existing literature (Berg *et al.*, 2008). None-

Table 11.3 Logistic regression result dependent variable: Y (correction of prediction)

Variable	*Description*	*Coefficient* (β_i)	*p-value*
	Constant	11.3044	0.9150
X_1	Number of contracts	−0.1970	0.0913
X_2	Number of active traders	−0.0171	0.0593
X_3	Closing trading volume	−0.0000	0.0940
X_4	Open interest	0.0000	0.0895
X_5	Number of traders with open interest	0.0164	0.1007
D_1	Dummy of Level_Local	−8.4837	0.9361
D_2	Dummy of Level_National	−7.2271	0.9456

Table 11.4 In-sample prediction results of the logistic equation

		Predicted category			
		0	*1*		
Actual	0	3	18	14.29%	Specificity
Category	1	1	150	99.34%	Sensitivity
				88.95%	Accuracy

theless, the signs of the number of active traders and closing trading volume are also negative, indicating that the more active the prediction market, the less likely it is that the prediction is correct. This finding is surprising from the aggregation perspective of the prediction market. Finally, open interest and the number of traders with open interest also have a positive effect. Based on our discussion in Section 2, if many traders still keep their positions open when the market closes, then they may believe that the closing price is not correctly set and great uncertainty over the market is expected, and hence it is less likely that the prediction will be correct. Therefore, the positive coefficients appearing here cast doubt on our hypothesis made for open interest.

4.2 Support vector machines

The logistic regression model is largely restricted to the assumption of a multivariate Gaussian distribution. From the histograms presented in Figures 11.1–11.3, and the basic statistics given in Table 11.2, the normality assumption fails to hold quite generally. In this situation, models which do not rest upon the assumption of normality will be attempted, and the support vector machine is the model which is frequently used in this situation.

The purpose is to use the suggested five variables to distinguish the successes from the failures.[14] These two sets are, however, not necessarily separable. What is normally done is to introduce non-negative slack variables, $\zeta_i \geq 0$, and a penalty, C, associated with these slacks, briefly, charging a cost for impreciseness. So we first introduce non-negative slack variables, $\zeta_i \geq 0$, to allow for the imprecise linear separation. From Equations (11.5) and (11.6), we have

$$\left(\langle w \cdot \varphi(x_i) \rangle - b\right) \geq 1 - \zeta_i \qquad \text{for } y_i = +1 \quad (\textit{success}) \tag{11.11}$$

and

$$\left(\langle w \cdot \varphi(x_i) \rangle - b\right) \leq -1 - \zeta_i \qquad \text{for } y_i = -1 \quad (\textit{failure}) \tag{11.12}$$

The optimal separation problem can then be expressed as a non-linear programming problem:

$$\min_{w,b,\zeta_i} \frac{1}{2} w^T w + C\sum_{i=1}^{N} \zeta_i \tag{11.13}$$

$$\text{subject to: } y_i\left(\left\langle w\cdot\varphi(x_i)\right\rangle - b\right) + \zeta_i \geq 0$$

$$\zeta_i \geq 0$$

Using the Lagrange multipliers and the Wolfe dual formulation, the problem is transformed to:

$$\max_{\alpha_i} L_{Dual} = \sum_{i=1}^{N} \alpha_i - \frac{1}{2}\sum_{i,j=1}^{N} \alpha_i\alpha_j y_i y_j K\left(x_i, x_j\right) \tag{11.14}$$

$$\text{subject to: } 0 \leq \alpha_i \leq C$$

$$\sum_i \alpha_i y_i = 0$$

To complete Equation (11.14), one has to choose a kernel function (Equation (11.8)). Common examples of the kernel function are the polynomial kernel, the hyper-tangent kernel and the Gaussian kernel. In this chapter, due to many of its mathematical virtues, we choose the Gaussian kernel:

$$\mathrm{K}\left(x_i, x_j\right) = \exp\left(-\left(2\sigma^2\right)^{-1}\left\|x_i, x_j\right\|\right), \tag{11.15}$$

where $\|, \|$ is the Euclidian norm. Equation (1.14) with Equation (11.15) can then be solved using the software LIBSVM.[15]

It is well-known that non-linear model tend to overfit.[16] Therefore, by following the standard practice, we divide the data set into the training set and the testing set. Two disjointed representative samples, each with 86 observations, are created. The testing result is given in Table 11.5. The accuracy rate is 86.05 percent, which is about the same as the logistic model, and is no better than simply guessing that all predictions are correct, which leads to an accuracy rate of 87.79 percent. The main problem again is that the obtained SVM is not able to identify the failures. In our testing set, there are 12 failures, but they are misclassified as successes; hence, specificity is 0 percent.

Table 11.5 Prediction results of the support vector machine

		Predicted category			
		0	*1*		
Actual	0	0	12	0.00%	Specificity
Category	1	0	74	100.00%	Sensitivity
				86.05%	Accuracy

4.3 Decision trees

A decision tree is also a model-free and data-driven tool. It has become the benchmark of machine learning. Unlike numeral networks or support vector machines, decision trees explicitly give a set of rules for making decisions, in our case, to decide whether the prediction made by XFuture tends to be right or wrong. To build decision trees, the software DTREG is applied.[17] The method used by DTREG to determine the optimal tree size is the well-known *K-fold cross validation.*[18] Using this method, one does not need a separate, independent data set for assessing the accuracy and size of the tree. Research has shown that little is gained by using more than ten partitions, so in DTREG ten folds are recommended as the default number of *K*. The decision tree derived based on the entropy maximization method and the ten-fold cross validation is depicted in Figure 11.4.

The tree started with the variable closing trading volume and distinguished the small trading volume (less than 21.5) from the rest. The starting rule is:

> If closing trading volume is "small" (≤21.5), then the prediction is correct.

This result is a little counter-intuitive, because it downplays the role of the activeness of the market, a result consistent with that of the logistic regression. The next decision is then based on open interest:

> Otherwise, if open interest is small (≤285), then the prediction is correct.

This rule favors the small open interest, which is the opposite to the result of logistic regression, but is consistent with our intuition of open interest. Continuing down the tree, we see the appearance of the non-monotone separation. Here, the decision variable is again trading volume, and the decision rule still disfavors the big trading volume.

> Otherwise, if both open interest and closing trading volume are large (>285 and >518,987), then the prediction is wrong.

However, the interplay of open interest and trading volume can enhance the predictions if the former is very large and the latter is large but not too large, as stated in the next rule.

> Otherwise, if open interest is very large (>97,675) but trading volume is neither not too large nor too small (between 21.5 and 518,987), then the prediction is correct.

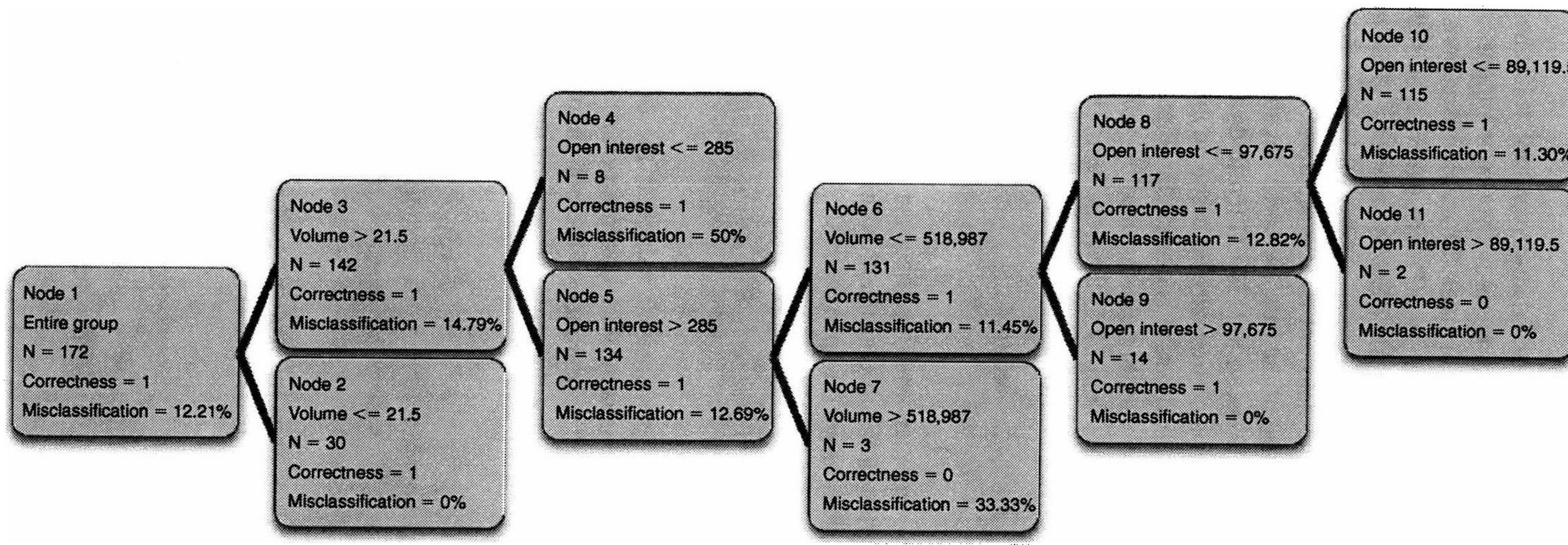


Figure 11.4 The derived decision tree.

Finally, to the very bottom of the tree, the last decision depends on open interest only. The rule is:

> Otherwise, if open interest is neither too small nor too large (285–89,119.5), the prediction is correct.

Hence, from the top to the bottom, the tree involves only two variables, namely, trading volume and open interest. The remaining four variables listed in Section 2.2 do not enter into any decisions here. The only two decision variables involved are both stratified, and the monotone (linear) decision rules are replaced by the non-monotone (non-linear) decisions with the respective stratification. The entire structure derived from the decision tree is, therefore, very different from the logistic regression.

The decision tree shown in Figure 11.4 was derived using a ten-fold cross validation, and the average of the ten-fold cross validation accuracy is shown in Table 11.6. Once again, the accuracy rate of 85.74 percent is not impressive enough to be comparable to the naive guess that all predictions are correct. It shares the same problem as the logistic regression and the support vector machine, namely, low specificity: only one out of 21 failures has been correctly identified.

4.4 Summary

To sum up, what is the Taiwan experience of prediction markets? An overall accuracy rate of 87.79 percent indicates that it works in Taiwan as well as it works in other places. Cultural factors might have influence on the micro details of the markets, but not much on the aggregate level.[19] As to what may cause the market to fail in some specific events, the answer is not clear at this stage. The attributes suggested by the literature have not helped us to distinguish the failures from the successes. At this point, no pattern has been found from the failures, as if they happened as accidents.

5 Concluding remarks

In this chapter we have made an attempt to investigate the prediction accuracy of the Exchange of Future Events (XFuture) in Taiwan. A data set composed of

Table 11.6 Prediction results of decision tree

		Predicted category			
		0	*1*		
Actual	0	1	20	4.76%	Specificity
Category	1	5	146	96.69%	Sensitivity
				85.47%	Accuracy

172 political events with 575 futures contracts from 2006 to 2010 is used in this study. Various quantitative models, including logistic regressions, support vector machines and decision trees, are applied to distinguish the correct predictions from the incorrect predictions. This study, therefore, becomes our first step toward opening the "black box" of prediction markets in light of empirical evidence. Our findings show that the various models perform quite closely, but the attributes which contribute to each model are heterogeneous, and are not all consistent with what we learn from the literature. Nonetheless, none of them are able to effectively distinguish failures from successes. One possible reason for this is that the sample size is too small and 21 failures are not enough to repeat themselves in any significant pattern. Hence, a study involving more events (as a total, we have 2,115 events) is definitely the next step. In addition to enlarging the sample size, there are a few other directions for further studies.

Design. First, in this chapter, we only consider the "winner-take-all" markets. As we said earlier, XFuture also runs share markets while the observations are still accumulating. One immediate issue is to see whether design matters as far as these two different types of markets are concerned.

Performance measure. Second, in this chapter, the prediction accuracy is simply based on the majority rule. Despite its simplicity, this rule does not allow for further distinction between 99 and 50.5 (or 49.9 and 0.1), i.e., between almost surely happen and barely happen. It would be useful to consider the alternative rules, which characterize the prediction error as a matter of degree rather than a dichotomous variable. It would then be interesting to see whether the six attributes listed above can better forecast these quantified errors.

Behavior. Third, in this chapter, not only is the number of traders included, but also *the type of traders.* The latter attempt is, however, just a beginning and is very primitive. It has not been conclusive regarding whether traders with open interest or some other characteristics can be a useful idea to examine the performance of prediction markets. However, we believe that it would be a fruitful direction to study the prediction markets from a *behavioral* aspect. The constituent types of market participants (the micro-structure) and the consequent effects on market performance have been studied quite intensively in the recent agent-based financial models (Chen *et al.*, 2010). In addition, the recent progress in behavioral economic experiments also sheds light on the contribution of the cognitive capacity and personality attributed to market participants (Chen and Wang, 2010). Corresponding to these studies, XFuture has already archived a database of individual traders. Hence, not only do we know the performance of each individual trader in each prediction market, but also his entire performance history. This information can help us to identify some *elite classes* of our traders, depending on the chosen criteria. The inquiry for the contribution of the elite class to the prediction market can then be analyzed, even though the Hayek hypothesis makes no explicit reference to the appearance of any elite class.

Notes

1 Regarding the early development of prediction markets in Japan, the interested reader is referred to Yamaguchi (2006). Also see http://shuugi.in.
2 www.ipredict.co.nz.
3 http://socioecono.phys.sinica.edu.tw/exchange/exchange_eng.html.
4 http://futures.nccu.edu.tw/exchange/exchange_eng.html.
5 Chen and Wu (2009) also addressed the pricing errors of the prediction market when its size is thin and provided several different solutions to correct for market illiquidity, from a simple median to complex wavelets. It is found that, after appropriate corrections, the prediction market can perform even better than the poll.
6 www.eel.nccu.edu.tw.
7 Wang *et al.* (2008) found that the topology of the trading network, which is hierarchical and scale-free, satisfied the famous scale-free network. On the other hand, they evidenced the existence of the Power law distribution in the asset return. These studies indicate that many familiar statistical properties generally existing in econophysics can also be found in the prediction markets.
8 www.xfuture.org.
9 http://pm.nccu.edu.tw.
10 XFuture also runs share markets, but due to space limitations, we have decided to focus on the "winner-takes-all" market in this chapter.
11 The way to define a correct prediction will be detailed in Section 2.
12 Of course, the market mechanism (the aggregation algorithm) also matters, but all of our 172 events are traded under the same design, namely, the "winner-takes-all" continuous-time order-book-driven double-auction market, which also supports shorting. This homogeneity, therefore, does not allow us to explore the role of designs in forecasting accuracy.
13 The closing price is the weighted average of all transaction prices on the final day of the market.
14 The two dummies are excluded because the support vector machine normally does not take qualitative variables, and that is the reason why we will try decision trees later in Section 4.3.
15 LIBSVM is developed by Chih-Jen Lin at National Taiwan University, and is freely downloadable at www.csie.ntu.edu.tw/~cjlin/libsvm.
16 In fact, using LIBSVM, we can perfectly separate the failures from the successes, i.e., a 100 percent accuracy rate.
17 www.dtreg.com.
18 In K-fold cross-validation, the original data set (in our case, 172 observations) is randomly partitioned into K subsets. Of the K subsets, a single subset is retained as the validation data for testing the model, and the remaining $K-1$ subsets are used as training data. The cross-validation process is then repeated K times, with each of the K subsets used exactly once as the validation data.
19 One issue related to the cross-cultural study is to examine whether the manipulation behavior of market participants differs among different countries' prediction markets.

References

Barr, J., Tassier, T., Ussher, L., LeBaron, B., Chen, S.-H. and Sunder, S. (2008), The future of agent based research in economics. *Eastern Economic Journal* 34: 550–565.

Berg, J., Forsythe, R., Nelson, F. and Rietz, T. (2008) Results from a dozen years of election futures markets research. In: Plott, C. and Smith, V. (eds.), *Handbook of Experimental Economics Results* Vol. 1, 742–751. Elsevier.

Chen, S.-H. and Tai, C.-C. (2006) On the selection of adaptive algorithms in ABM: a computational equivalence approach. *Computational Economics* 28(1): 51–69.

Chen, S.-H. and Wu, W.-S. (2009) Price errors from thin markets and their corrections: studies based on Taiwan's political futures markets. *Advances in Econometrics* 24: 1–25.

Chen, S.-H. and Hsieh, Y.-L. (2010) Reinforcement learning in experimental asset markets. *Eastern Economic Journal*, forthcoming.

Chen, S.-H. and Wang, S.G. (2010) Emergent complexity in agent-based computational economics. *Journal of Economic Surveys*, forthcoming.

Chen, S.-H., Chang, C.-L. and Du, Y.-R. (2010) Agent-based economic models and econometrics. *Knowledge Engineering Review*, forthcoming.

Grunwald, P. (2007) *The Minimum Description Length Principle*. MIT Press.

Quinlan, R. (1986) Induction of decision trees. *Machine Learning* 1(1): 81–106.

Quinlan, R. (1987) Simplifying decision trees. *International Journal of Man-Machine Studies* 27(3): 221–234.

Quinlan, R. (1993) *C4.5: Programs for Machine Learning*. Morgan Kaufmann.

Smith, V. (1982) Markets as economizers of information: experimental examination of the "Hayek Hypothesis." *Economic Inquiry* 20(2): 165–179.

Tseng, J.-J., Chen, S.-H., Wang, S.-C. and Li, S.-P. (2008) Scale-free network emerged in the markets: human traders versus zero-intelligence traders. In: Terano, T., Kita, H., Takahashi, S. and Deguchi, H. (eds.), *Agent-Based Approaches in Economic and Social Complex Systems*, Vol. 6, 245–254. Springer.

Tseng, J.-J., Chen, S.-H., Wang, S.-C. and Li, S.-P. (2009) Emergence of scale-free networks in markets. *Advances in Complex Systems* 12(1): 87–97.

Tung, C.-Y., Chou, T.-C., Lin, J.-W., Lin, H.-Y. and Liu, C.-K. (2009) An analysis on the accuracy of election forecasts: case studies of 2006 and 2008 elections in Taiwan. Working Paper, Center for Prediction Markets, National Chengchi University.

Vapnik, V. (1998) *Statistical Learning Theory*. John Wiley.

Wang, S.-C., Liu, K.-P., Yu, C.-Y. and Li, S.-P. (2004) A web-based political exchange for election outcome predictions. *Proceedings of IEEE/WIC/ACM International Conference on Web Intelligence*, 173–178.

Wang, S.-C., Tseng, J.-J, Li, S.-P. and Chen, S.-H. (2006) Prediction of bird flu (H5N1) outbreaks in Taiwan by online auction: experimental results. *New Mathematics and Natural Computation* 2: 271–279.

Wang, S.-C., Tseng, J.-J., Tai, C.-C., Lai, K.-H., Wu, W.-S., Chen, S.-H. and Li, S.-P. (2008) Network topology of an experimental futures exchange. *European Physics Journal B* 62: 105–111.

Wang, S.-C., Li, S.-P., Tai, C.-C. and Chen, S.-H. (2009) Statistical properties of an experimental political futures market. *Quantitative Finance* 9(1): 9–16.

Wolfers, J. and Zitzewitz, E. (2004) Prediction markets. *Journal of Economic Perspectives* 18: 107–126.

Yamaguchi, H. (2006) General Election Hatena: the first political prediction market in Japan. *Proceedings of the 6th Computational Intelligence in Economics and Finance*. Atlantis Press.

12 Uses of sports wagering-based prediction markets outside of the world of gambling

Rodney J. Paul and Andrew P. Weinbach

Gambling markets have provided a fertile ground to study the efficient markets hypothesis. The use of prices in these markets – such as odds, point-spreads and totals – have served to test whether all available information is included in the current prices formed in these wagering markets. Given the simple nature of this market, including a clearly defined starting and end point, quick realization of returns and usual ample liquidity, the fundamental notions behind the efficient markets hypothesis are generally straightforward to test.

In general, betting markets have offered support in favor of the efficient markets hypothesis, especially in the market for all games in a given sport over a long time horizon. In a market dominated by staunchly loyal fans and sometimes crazed enthusiasm, these results give a stamp of approval to the notion of efficiency and support the logic behind the wisdom of crowds. Along the way, in various subsets of betting market data, phenomenon such as the favorite–longshot bias (and its reverse) and behavioral biases such as preferences for good teams (manifested as road favorites and big home favorites) and scoring (more preferred to less) have led to some interesting questions in behavioral finance about how investors and consumers truly behave.

One area of this study that is often overlooked – given the importance of the efficient markets hypothesis as it relates to fundamental theories in finance – is how prices formed in betting markets may have outside uses in economics and other areas related to business. Instead of being the sole proprietorship of finance research, point-spreads, odds and totals are now being linked to other areas of research. What was once thought of as only a study of gamblers for the ease of use of studying market efficiency is now finding new outlets where this information can be quite valuable to researchers and people working in the sports business world.

These outside applications of gambling market data may not seem straightforward at first, but the usefulness of the information provided by these markets stems from what we understand about the concept of prices. Free and open markets do an incredible job of providing information. Wants and needs of consumers, desires and cost structures of producers, and expectations of all parties are amalgamated into a single number. Prices formed in financial markets, such as betting markets for sports, are assumed to include all available information at the current moment. Prices in financial markets are expected to provide an optimal and unbiased forecast of future events; in the case of gambling markets, this future event is the outcome of the game. Even in cases where behavioral biases may result in a rejection of the efficient markets hypothesis, this too provides researchers with important information, because systematic deviations from unbiased forecasts may in fact be revealing more information about what bettors, and likely fans, enjoy about their particular sports and teams.

Having some foresight into the outcome of a sporting event in terms of which team is likely to win and by how much is important information to many audiences. Obviously, fans of the teams care about this in the days and hours they spend thinking about the upcoming game for their team, and gamblers care about these prices as they attempt to place a winning wager. This information spreads well beyond these groups, however, as the teams themselves likely care a great deal about the expectations of their fans as it could easily affect ticket sales, concession sales (i.e., fans leave early in a blowout) and merchandise sales. Television networks are likely to find this information useful in deciding which games to broadcast. Large television entities such as ABC/ESPN have a number of networks where they broadcast many different games, often from the same league and sport. Having some insight into the most enticing match-ups of the day or week can be very important to their bottom line. Similarly, advertisers wanting to achieve a maximum return for their advertising dollar would like to place their ads where the most viewers will be watching. Prediction markets in the form of gambling markets provide this information and likely much more.

Some lines of investigation using prices formed in gambling-related prediction markets to research a concept other than market efficiency have a substantial history, while others are just beginning to emerge. What we wish to illustrate in this chapter is some of the lines of research where these betting market prices are being used to answer other questions in economics in general and, more specifically, in sports-related businesses. To do this, we have decided to discuss past research in a few distinct areas that use odds, point-spreads, totals or other wagering-related prices as sources of useful information.

One natural course of study has been in the realm of the uncertainty-of-outcome hypothesis. If fans prefer games to have uncertain outcomes, this should be reflected in things such as attendance and television ratings. We outline and describe the research in this area along a few different lines. We first describe tests of the adequacy of betting market formed prices in being a proxy for uncertainty of outcome and its applications to the sport itself. We then outline the background and specific tests performed in past research to determine if prices formed in betting markets have any impact on attendance. We also investigate the role that wagering market prices play in relation to television ratings. Lastly, we describe some research that would fall into the "other" category, where these prices are used to investigate other areas of the sports world beyond the uncertainty-of-outcome hypothesis. Each section presents a summary of past research on these topics, with some natural overlap of categories for studies which pursued multiple conceptual goals.

1 Uncertainty of outcome in general sports analysis

Before discussing the main use of the uncertainty-of-outcome hypothesis in relation to its effects on attendance, it is useful to note some recent studies which use betting market prices to study uncertainty of outcome in a more general manner (not directly related to attendance). The main question concerning uncertainty of outcome is how to effectively measure it? It appears to be a different concept from competitive balance, which is closely related, but typically calculated in an *ex post* fashion. To measure uncertainty of outcome in a literal sense, it would appear that *ex ante* measures would be appropriate, and researchers have explored various approaches to try to capture the *ex ante* expectations of fans.

A variety of possibilities exist to estimate uncertainty of outcome, including win–loss records from the current and previous seasons, game statistics representing various match-ups, and other possible historical measures. An alternative to purely past data lies in the realm of prediction markets. The gambling market for sports serves as an excellent prediction market of future events. Therefore, prices formed in prediction markets, such as odds and point-spreads, may serve as an excellent measure of uncertainty of outcome in a given sports league.

The goals in measuring uncertainty of outcome are quite similar to the goals in measuring competitive balance. The researcher in this field is striving to come up with the best measure of how evenly matched teams are in a given game or over the course of a season. Competitive balance is measured after the games are played, which could still be quite useful for predicting future decisions by fans to attend or watch games, but uncertainty of outcome is measured for each game, before it is played, through a price formed in a prediction market.

The odds on a game reflect how likely the favorite is to defeat the underdog. The closer to even money a game is, the more likely the game is to be close. The larger the favorite odds, the more likely the game is to be a blowout. Assuming fans prefer to watch close games as opposed to lopsided contests, the betting market odds can be quite useful as a proxy for this important factor in the minds of fans. Uncertainty of outcome also offers information that competitive balance figures cannot hope to capture, such as injuries, suspensions and weather, which could all be important to a contest, but would not be reflected fully in *ex post* information such as win percentages of teams.

One recent example where betting market odds were used was for Major League Baseball. Betting market odds on baseball games were used to illustrate the importance of uncertainty of outcome in the general perceptions of the fans and media as they relate to this sport. For a sample of years covering 1990–2006 for Major League Baseball, Paul *et al.* (2009) used the average of betting market odds as an *ex ante* measure of uncertainty of outcome to compare to competitive balance measures in the National and American Leagues in Major League Baseball.

The results of this study were compared to the findings of Schmidt and Berri (2001), who had previously shown that the 1990s was the most competitive decade in the history of Major League Baseball. Schmidt and Berri (2001) studied competitive balance, an *ex post* figure, using a measure based on the win percentages of the teams. Although competitive balance studies are quite useful and have been performed across sports, there is reason to believe that an *ex ante* measure of competitiveness of a certain game, its uncertainty of outcome, could be more valuable than an *ex post* figure. Given that fans make the decision to attend games and pay for tickets or take their time to watch a game on television, knowing something about the likely closeness of a game before it is played could be quite valuable to teams, leagues and advertisers.

For baseball, average favorite odds on both American League games and National League games were shown to steadily increase during the 1990s. This uncertainty of outcome measure increased by more than 10 percent, from the low 130s on the favorite at the start of the decade to a level above 150 by the end of the decade. These odds spiked for both leagues even further in the early 2000s. Comparing this to the standard deviation of win percentage and GINI coefficients for the leagues (measures of competitive balance) during this time-frame reveals that while the measures of competitive balance remained mostly stable, the uncertainty of outcome figure (favorite odds) steadily increased.

These results reveal that bettors, and likely fans, believed that outcomes of baseball games were becoming less competitive during the 1990s. This uncertainty of outcome result could help to explain why fans and sportswriters complained of a lack of "competitive balance" during the 1990s, even though the true standard *ex post* competitive balance figures did not bear this out. In truth, fans and sportswriters may have actually meant the uncertainty of outcome of games decreased during the 1990s, meaning that baseball games appeared to be less interesting as more games contained a more dominant team versus a lesser team in the contest. This prediction market-formed price, betting market odds, therefore can be quite useful in explaining how fans, sportswriters and others may view upcoming baseball games.

Another sport where uncertainty of outcome was studied using prices formed in prediction markets was for NCAA Football. College sports in the United States has a schedule structure where teams play most, if not all, teams within their conference and play a few out-of-conference games. Some teams schedule very tough opponents outside of the conference, while other schools schedule quite easy games. Given this imbalance in the schedule, win percentages and competitive balance figures can be quite unreliable, as a team with a lower win percentage could quite easily be a far superior team to one with a higher win percentage, based on the quality of their opponents.

With this in mind, one way to tackle the problem of measuring the quality of teams and the quality of match-ups between teams is through the use of point-spreads formed in the college football betting market. Given overall market efficiency in the college football wagering market, the point-spread should offer an optimal and unbiased forecast of the outcome of a game, with its key advantage being that it is known before the game is played. Given that this *ex ante* information is likely quite important to fans deciding to buy tickets or watch games, voters in the polls deciding upon the relative quality of teams, colleges and conferences attempting to maximize profits, etc., the average point-spread in college football games is likely quite meaningful to study.

One pertinent point of study for college football is the effects of the introduction of the BCS (Bowl Championship Series). College football determines a champion though a series of polls and computer rankings leading to the top teams in the BCS rankings being placed in the prime bowl games, with the top two teams pitted against each other in the BCS championship game. Given the subjective rankings of those participating in the polls and the factors which affect computer rankings, perceived quality of teams is very important to the financial bottom line of these major colleges. Although it might be reasonably straightforward to compare teams within-conference, comparing out-of-conference teams tends to be quite difficult.

A key element of the BCS system is at-large bids to the BCS bowls and the big financial payoff these bids ensure. While the champion of a major (BCS) conference gets an automatic bid to these prime bowl games, the conference can also get a second at-large bid if the overall quality of the second team in the conference is higher than in other conferences. In addition, non-BCS conference teams can achieve a BCS bid if they win their conference and rank in the top 12 of the overall end-of-season BCS poll. These rules provide different incentives for the BCS conference teams compared to the non-BCS conference teams. Conferences in the BCS would prefer the perception (and likely reality) of having a strong conference as this makes the valuable second at-large bid more likely. Non-BCS conferences, on the other hand, need a dominant (likely undefeated) team to have a chance at cashing in on BCS glory.

Given these possible incentives, Paul *et al.* (2010) tested the level of uncertainty of outcome in BCS and non-BCS conferences both before and after the initiation of the BCS system. Using the average favorite point-spread as the measure of uncertainty of outcome, the authors tested whether BCS conferences have greater uncertainty of outcome in their games after the introduction of the BCS and its at-large bid incentives and if non-BCS conferences have less uncertainty of outcome, due to the benefits for the conference created by producing an undefeated team.

Using data from 1985 to 2008, it is shown that four of the six BCS conferences had a statistically significant decrease in the average point-spread after the introduction of the BCS. This means that in four of the conferences (ACC, Big 10, Big 12 and SEC) uncertainty of outcome was increased in conference games after the introduction of the new system. In the non-BCS conferences studied (MAC, WAC, Conference USA and Mountain West), all conferences showed an increase in the average point-spread (less uncertainty of outcome) with the introduction of the BCS.

These results suggest that the conferences may be responding to the incentives of the at-large bid system of the BCS, as one would expect. Big conferences became more competitive (more uncertainty of outcome) due to possibilities of a highly respected conference being more likely to attract an at-large bid. Non-BCS conferences, on the other hand, became less competitive as their only real chance at an at-large bid to the BCS lies with an undefeated team.

2 Betting market prices, uncertainty of outcome and attendance

The primary use of the uncertainty-of-outcome hypothesis by researchers has been to test for effects on game attendance. The basic premise behind this idea is that if fans enjoy uncertainty of outcome, in that they enjoy close games between relatively evenly matched teams, measures of uncertainty of outcome should have a significant effect on attendance. Odds and point-spreads from betting markets are used as measures of uncertainty of outcome and are included in regression models of attendance.

For those unfamiliar with these studies, it is important to note that the wording of these results in relation to uncertainty of outcome can become quite tricky. For instance, consider a decrease in the average odds on league matches. This represents an increase in the uncertainty of outcome as the games are expected to be closer. This increase in the uncertainty of outcome should lead to an increase in attendance, if this is what fans prefer. An increase in average odds, on the other hand, leads to less uncertainty of outcome as favorites are becoming more dominant in these games. This increase in odds should lead to a negative effect on attendance, if fans prefer uncertainty of outcome.

The effects of uncertainty of outcome as it pertains to game attendance have been studied for a variety of sports across the globe. To group these studies and discuss their results, we will start with the non-North American sports and then discuss what is seen in the major sports in North America. There have been multiple attempts to determine whether uncertainty of outcome drives attendance in the English football (soccer) league. Peel and Thomas (1988) consider betting odds as a good *ex ante* measure of uncertainty of outcome, suggesting that departures from efficiency in this market appear to be small. The authors examine attendances in the English football league for the 1981–1982 season, and conclude that fans appear to enjoy uncertainty of outcome, as proxied by betting odds, but also prefer to attend games with good teams. The authors caution that attempts to alter league policies to make team ability more equal may be harmful, since doing so would reduce the extent to which certain teams are "good," which fans also find appealing.

Peel and Thomas (1992) also studied English football in the 1986–1987 season, concluding that fans do not wish to attend games with very high odds on the favorite. In addition, Peel and Thomas (1992) found that fans enjoy it when the home team is favored, meaning that, all else being equal, fans would prefer to see a close game in which the home team prevails. Expected goals scored, proxied by the betting market over/under (total) was also shown to have positive and significant effects on attendance. This illustrates a clear preference for more scoring compared to less.

Forrest and Simmons (2002) also studied English football. They researched the 1997–1998 season and found support that uncertainty of outcome helps to increase attendance. Using betting market odds and correcting for behavioral biases in the market, Forrest and Simmons (2002) find that fewer fans attend games when the adjusted odds on the favorite are quite high, while more fans attend games as the odds on the favorite fall.

All three divisions of the Scottish football league were studied by Peel and Thomas (1996) for the 1991–1992 season using betting market odds as a measure of uncertainty of outcome. Based on the model of Theil (1967), maximum uncertainty of outcome is found to occur when the probabilities of a home win, road win and draw are all equal. The authors find a U-shaped relationship between attendance and uncertainty of outcome and conclude that fans do not have a preference for uncertainty of outcome in Scottish football.

Buraimo and Simmons (2009) also show that fan attendance is not influenced by uncertainty of outcome in the traditional expected manner. In a study of the Spanish Primera division, fans attending games prefer less uncertainty of outcome in Spanish football. Using betting market odds to construct the probability of a home win, Buraimo and Simmons (2009) show that attendance is maximized when there is a high home team win probability or there is a low home team win probability.

By using the home team win probability and its square within the attendance regression, the authors show there are two effects at work. Fans like to see the home team win and therefore are more likely to attend games when the home team is a large favorite. In addition, fans prefer games with low home team win probabilities. This may be confusing at first, but the authors explain this as a "David vs. Goliath" effect. Fans prefer to attend games where the home team may upset a powerful opponent as this will give people who attend the game "bragging rights" about being there for the great upset. The quality of the opponent may play a key role here as well, as individual dummy variables for perennial powers in this league such as Barcelona and Real Madrid were shown to have positive impacts on attendance on the road.

Overall, these results concerning attendance reject the uncertainty of outcome hypothesis, as fans do not appear to enjoy expected close matches in the Spanish Primera division. Buraimo and Simmons (2009) also study the television audiences for Spanish football, which is further discussed later in this chapter, along with other findings related to television viewing audiences.

Rugby has also attracted considerable attention in terms of the testing of uncertainty of outcome on attendance using betting market odds. In a study of British Rugby League matches in the 1994–1995 season, Peel and Thomas (1996) determine that outcome uncertainty plays a significant role in the determination of attendance for rugby. Fans of rugby appear to enjoy uncertainty of outcome as betting market odds on this sport have a negative and significant effect on attendance. Therefore, as the odds on the favorite increase, fewer fans attend these games.

Carmichael *et al.* (1999) also found statistical evidence of the uncertainty of outcome being important as it relates to rugby attendance. Using the 1994–1995 season of English Rugby League fixtures, the authors find that attendance is lower for games with longer match odds. This means that as the odds on the favorite increase, fewer fans attend rugby matches. Again, this suggests that fans do enjoy uncertainty of outcome in this sport. In addition, Carmichael *et al.* (1999) also examined the pre-season odds of each team winning their division and found that teams which were longshots to win their divisions had lower attendance, while teams which were favorites or had shorter odds to win the division had higher attendance at their matches. This result also supports the notion of the importance of uncertainty of outcome as fans prefer to watch teams that have a chance to win their division and overall title.

For rugby in New Zealand, however, uncertainty of outcome was not shown to have an impact on attendance. In a study of the Super 12 rugby league for the 1999–2001 seasons, Owen and Weatherston (2004) used betting odds as a proxy for uncertainty of outcome. Based on attendance at individual matches, the authors find little evidence that uncertainty of outcome has any effect on attendance.

Uncertainty of outcome in cricket matches was studied by Morley and Thomas (2007). In an examination of the limited overs cricket matches for 1996 and 1997, the authors found that uncertainty of outcome apparently was not preferred by fans. As the odds on the favored team increased, more fans actually attended the games.

For research examining fan interest based on uncertainty of outcome (as measured by betting odds) for soccer, rugby and cricket, the results are quite mixed. In some cases, the theory holds true as higher match odds on the favorite result in lower attendance. In other cases, no effect or the exact opposite is found. These results call into question a universal preference for uncertainty of outcome across these sports and deserve further study to determine if uncertainty of outcome is being measured correctly and/or what aspects of these sports may cause fan demand to differ.

In the major North American sports, the relationship between attendance and uncertainty of outcome has also been studied and the results have generally been consistent. Welki and Zlatoper (1999) examined the National Football League. They used actual game attendance as a proportion of total tickets sold as their dependent variable to capture both the number of people in attendance and those with tickets who chose not to attend the game. The point-spread and point-spread squared were used as proxies for uncertainty of outcome in NFL games and found support for the uncertainty-of-outcome hypothesis. Games with smaller point-spreads, all else being equal, were shown to generate higher attendance.

In Major League Baseball, Knowles *et al.* (1992) studied the 1988 National League season. They used betting market odds to construct the probability of a home team win, which served as a proxy for uncertainty of outcome within the games. The authors found that uncertainty of outcome, as measured through the odds, was a significant determinant of attendance. This led to their conclusion that the maximization of fan attendance would occur when the home team was a slight favorite. Rascher (1999) studied the 1996 baseball season and found similar fan preference for uncertainty of outcome. In general, large favorites were not found to be popular with the ticket-buying baseball audience. Like Knowles *et al.* (1992), slight home favorites outperformed other favorites at the gate.

Recently, Lemke *et al.* (2010) found the opposite result for the 2007 Major League Baseball season. The betting market odds on baseball games were included in an attendance model and used as a measure of expected win probability, reflecting uncertainty of outcome. This variable was included alongside many other explanatory variables, many directly correlated with the betting market odds, and the opposite result compared to the other baseball studies was found. Lemke *et al.* (2010) found that fans in the 2007 season preferred a lack of uncertainty of outcome as attendance increased as the probability of the home team winning increased. This result could represent a fundamental change in the way baseball fans viewed their sport, or the relationship between the betting market odds and other explanatory variables may confound this result.

Potential problems with using betting market odds as a measure of uncertainty of outcome were noted by Dawson and Downward (2005) in the *Journal of Sports Economics*. Dawson and Downward (2005) note two potential difficulties in using betting markets in this capacity. First, they note that there could be biases in betting markets, which lead prices formed in these markets to not truly reflect an optimal and unbiased forecast of the outcome of a game. More importantly, in their opinion, they believe that uncertainty of outcome is inherently a quadratic relationship. Therefore, they explain that betting market odds are actually a measure of the probability of a home team win, not true uncertainty of outcome. Inclusion of the quadratic term related to betting market prices might be helpful with this problem. The authors note, as many do in relation to the economic concept of comparative advantage, that they do not seek to declare one method better than others in relation to uncertainty of outcome, but are exploring various methods in hopes of finding the best measure possible.

While the results are mixed across different leagues of the same sport and different sports in general, there is at least some support for the theory that the prices formed in prediction markets may signal what fans truly prefer in a sporting event. The important research question that remains to be fully answered is: Are studies that incorporate betting market odds and point-spreads using an improper measure of uncertainty of outcome, leading some studies to detect no (or the opposite) effect of these variables, or is it incorrect to assume that fans actually value uncertainty of outcome when it comes to sports? If we assume that fans do enjoy uncertainty of outcome, then further work needs to be done to refine our odds measure of uncertainty of outcome to truly represent this variable. Some possibilities include modeling uncertainty of outcome as a quadratic relationship and testing the difference between the effects of using odds at market open compared to market close.

On the other hand, if fans do not actually prefer uncertainty of outcome in sports, then our fundamental assumption is wrong and it may not be the prediction market-based pricing which is inaccurate. Another difficulty in estimating the effects of uncertainty of outcome is that it certainly may not be homogeneous across sports. In any event, the relationship between uncertainty of outcome and attendance appear fruitful ground for future research as sports economists attempt to better understand what truly influences fan behavior.

3 Betting market prices, uncertainty of outcome and television ratings

In a study of the English Premier League, Forrest *et al.* (2005) use the uncertainty-of-outcome hypothesis to study television viewership. The advantages of studying television viewers as opposed to attendance at games, as the authors note, is that the television audience is much larger, researchers studying attendance at stadiums must deal with the possibilities of sell outs, and the game audience in attendance at the stadium is usually mostly driven by home team fans. In modeling television viewership, the expected closeness of the game may be much more observable and important.

Data for this study was gathered from SKY Broadcasting, which broadcast 60 (later 66) matches from the English Premier League. Their sample includes the 1993–1994 season through the 2001–2002 season. Games on television were played on Sundays and Mondays to avoid any negative effect the televised contest may have on actual attendance in-person at stadiums on Saturday. Forrest *et al.* (2005) note the broadcasting differences for matches which occur pre-Boxing Day to post-Boxing Day, as the pre-Boxing Day schedule is set at the start of the season, while flexible scheduling becomes available after Boxing Day. Presumably, the flexible scheduling portion of the season would take into account uncertainty of outcome in the decision-making of the network.

The authors discuss the use of betting market odds to measure uncertainty of outcome, but ultimately decide against its use due to possible biases in the betting market data. Although market efficiency studies of the English Premier League, other soccer leagues and other sports in general do not reveal vast profits for bettors, as markets are generally found to be efficient, some biases such as a reverse favorite–longshot bias in odds-based sports and a clear preference for favorites (road favorites especially) tends to exist in point-spread markets. These biases are generally not enough to earn profits, but they could allow for odds that do not truly reflect the actual uncertainty of outcome of a game.

With this in mind, however, it is important to note that the biases of bettors (either known by the sportsbook and priced accordingly or evident through the flow of bets and subsequent changes in odds) may not truly be different from the biased notions of fans in general. Fans may very well have the same (if not more) slightly skewed beliefs of who will actually win a contest. Bettors have financial incentives to estimate the outcome of a match in an unbiased manner, while fans do not. In this financial market, bettors who constantly have biased perceptions of a team or teams will be punished financially and may learn from their mistakes. Fans, on the other hand, do not take financial losses over time due to their biased perceptions.

We believe that any behavioral biases that exist in gambling markets likely provide information, rather than subtract from its use. We believe that the groups of fans and bettors have a large overlap and that biases present in one group are likely to be present in the other. Therefore, if a team is constantly overbet, leading to inflated prices in betting markets, they are likely to be quite popular at the gate and on television, likely in excess of teams which do not show these inflated prices. Although the measure of uncertainty of outcome used in Forrest *et al.* (2005) are quite plausible through the use of win percentage differences and a home field advantage based on previous season results, it cannot hope to assimilate all of the information (including bettor biases) which the odds, as a single price in a financial market, can capture.

Forrest *et al.* (2005) found interesting and important results in their study of uncertainty of outcome as it was shown to affect the choice of games and the number of viewers of a game (post-Boxing Day) in the expected manner. It would be quite interesting to use odds (and odds squared) as an alternative measure of uncertainty of outcome, both to see its effect in the current model used by the authors, but also to see its interaction with some of the other independent variables used in this study. The authors account for variables such as wages in the cities where the teams are located and dummies for historically successful or popular teams. Perhaps the betting market odds, which may include some biases toward the most popular teams and biggest cities, may also capture these effects in the one variable. In any case, this would be informative to see side-by-side, and this and other similar studies may push our knowledge forward on uncertainty of outcome in the future.

In a study of the effect of betting market odds on attendance and television ratings in Spanish football, Buraimo and Simmons (2009) found quite different results for each set of spectators. As mentioned in the previous section, their results for the Spanish Football Primera division revealed that fans prefer home favorites with a high win probability (heavy home favorite odds) and also prefer when the home team is a substantial underdog (heavy road favorite odds). However, the authors do not find evidence of comparable preferences among television viewers.

Using two-stage least squares, Buraimo and Simmons (2009) show that the absolute probability difference between the teams (difference between home and away win probabilities based on betting market odds) has a negative and significant effect on the size of the television audience for Spanish football. As this difference gets larger, meaning that one team is becoming more of a prohibitive favorite in the match, the number of viewers on television falls by a significant margin. This result suggests that fans of this league prefer uncertainty of outcome in matches on television.

The apparent differences in preferences between those who attend live matches and those who watch on television may be driven by differing preferences for the home team in each setting. In most cases, fans turn out at the stadium to watch the home team play and generally have a rooting interest in that squad. For nationally televised games, however, there would seem to be little cause for such bias for the home team. In nationally televised games, there could very well be more of a fan interest in one team rather than the other, but this is not dependent upon which team is the home team. Regional television coverage may have similar results to attendance models, whereas neutral-site games, perhaps as part of mega-events, may have similar fan preferences to the television results shown in Buraimo and Simmons (2009).

The role of prediction market-formed prices was also used in studies of sports in the United States, although in a slightly different manner than some of the European studies. In the National Football League, historically one of the biggest games of the week is the Monday Night Football contest. This game is shown on ESPN, a cable network that is available to most US households, and is the sole NFL broadcast of the night, and has typically been a big ticket item for the NFL, with ESPN currently paying $1.1 billion annually for broadcast rights. In recent years, with the addition of flexible scheduling for the final weeks of the NFL season, choices about which game to put on in this time slot have changed from a pre-season decision to a within-season decision, which hopefully allows for more meaningful match-ups and higher ratings.

Nielsen ratings for Monday Night Football were studied for the years from 1991 to 2002 to determine if viewers prefer uncertainty of outcome and scoring. During this timeframe, the Monday Night Football game was broadcast on ABC television, but moved to ESPN (an ABC/Disney property) after the timeframe studied in this sample. The goal of the study was to identify factors, both before the game was played and during the game itself, that determine the television ratings for NFL Monday Night Football games.

In this context, prices formed in prediction markets (specifically, the betting market for the NFL) may be useful in identifying and improving our understanding of fan interest. Although the point-spread is somewhat problematic to study in this context, due to a small point-spread either reflecting two high-quality teams or two low-quality teams competing against each other, inclusion of the total in this model provides some potentially helpful information which might otherwise not be available.

The Las Vegas total posted on an NFL game represents the number at which bettors can wager on two simple propositions, the over or the under. An over bettor will win if the combined score of both teams in the game exceeds the posted total, while an under bettor will win if the combined score of both teams is less than the posted total. In an efficient market, which the NFL totals market has been shown to be for the sample as a whole, this price should represent an optimal and unbiased forecast of the number of points which will be scored in the game.

The simple premise of this study by Paul and Weinbach (2007) in the *Journal of Economics and Business* is that fans prefer higher-scoring games to lower-scoring games. In addition, they prefer to watch games which are *expected* to be higher scoring. Therefore, the Las Vegas total should serve as a good proxy for expected scoring and, therefore, fan interest. If higher-scoring games are truly more exciting than lower-scoring games and television viewers (fans) prefer this, the totals should have a positive and significant effect on Nielsen ratings. Games that are expected to be higher-scoring should have higher initial ratings for

Monday Night Football, compared to games that are expected to be lower-scoring, all else being equal (winning percentages of the teams, point in the NFL season, etc.).

The beginning-of-game (9:00 p.m. EST – during this timeframe – the start of the Monday Night Football game) Nielsen ratings were used as the dependent variable in the first regression model in this paper. The independent variables included a variety of time-specific variables (months of the year, yearly dummies, dummy for the World Series (baseball) being on opposite Monday Night Football), variables for the quality of the teams in the match-up (difference in win percentage to account for uncertainty of outcome and sum of win percentage to account for overall game quality), and the Las Vegas total to proxy for expected scoring.

Fans were found to enjoy uncertainty of outcome, as more evenly matched teams received higher-rated games, and enjoyed seeing contests between the best teams in the league as proxied by the sum of the win percentages. Dummy variables for various time-related variables had the expected results and were generally significant as anticipated. The total variable, the key expectations variable related to prediction markets in this study, was found to have a positive and significant effect on the initial ratings for Monday Night Football. This means that more viewers tuned in at the start of the game if the game was expected to be high-scoring. This supports the notion that fans enjoy more scoring compared to less.

Although not directly related to prediction markets, the study also found that fans are more likely to turn the game off at halftime if the score differential is not close and if there is not much scoring. The total number of points scored at halftime was found to have a positive and significant effect on the Nielsen ratings during the 10:30 p.m. time slot on ABC.

This example demonstrates the potential value of prediction markets in quantifying fan expectations of upcoming games, as well as fan perceptions of attributes of individual sports teams, including their quality and expected scoring ability. This is a clear example of how prices formed in financial markets can be a useful tool to those outside the world of gambling. In this case, the market provides information about a widely consumed product, televised sports, before a game is played. As prices produced by these markets improve our understanding of fan preferences, professional sports leagues, networks and advertisers could potentially gain substantial benefit from this information. Potential applications include improvement of league rules, including revenue sharing and scheduling policies, further optimization of game broadcast patterns and improved results for advertisers during these games.

4 Sports prediction markets and voting in college sports polls

Another example of where prices formed in sports gambling markets can be useful outside the world of gambling occurs in the polling process used in college football. College football in the United States has a rather unique way of determining a champion. In order to get to play in the national title game, teams must be at the top of the BCS rankings. These rankings are determined partly by computer, but also through the use of polls. Polls reflect the opinions of those deemed experts on college football and are often met with much skepticism and, in some cases, even contempt by fans, players and coaches.

Much of the reason for the controversy around polls results from a difficulty in explaining how and why voters choose teams in the order that they do. Winning and losing are obviously one method of making this choice, but in a college football setting with many conferences, divisions and teams, and many top teams never playing each other in the regular season, there is obviously more that goes into the decision-making than simply wins and losses.

Point-spreads, determined by betting markets, may be helpful in improving our understanding of the college football polls as well. Using information on teams and voters in the two important polls (at the time) in the BCS rankings, the ESPN/USA Today coaches' poll and the AP writers' poll, for the 2003–2004 season, Paul *et al.* (2007) introduce the point-spread into a model of voter behavior. Specifically, the point-spread is introduced as a proxy for expectations of voters for particular games. Given that the point-spread market has been shown to be overall efficient in previous studies of college football, the point-spread should represent an optimal unbiased forecast of the outcome of the game. The point-spread of a particular game becomes a useful point of reference against which to compare the actual game outcome.

If the score differential in the game exceeds the posted point-spread, the favorite team has won by more than expected. If the price formed in this financial betting market does serve as a good proxy for expectations, the favorite team in this case should be rewarded. On the other hand, if a favored team loses a game outright, or perhaps more importantly to this study, if a favored team wins, but does not cover the posted point-spread, they are likely to be downgraded in the polls. Since we assume that voters in both the coaches' and writers' polls are ranking teams based upon their apparent strengths and weaknesses, new information which becomes available through actual performance differing from expected performance should be included in their relative rankings.

Of course, it is possible that different coaches and writers have quite different expectations for teams, but it is likely the point-spread serves as a good proxy for the average expectations within this market. Given the time demands on coaches and writers, it is also unlikely that these voters actually watch each game that every college football team plays. Therefore, the posted point-spread may provide information to these voters, if they do not have strong feelings one way or the other about the strengths and weaknesses of the teams involved in the game, and a simple comparison of the actual game outcome compared to the expected game outcome may influence their weekly rankings of teams.

Paul *et al.* (2007) construct a regression model to study the role of the point-spread as a proxy for expectations of college football voters. The dependent variable in the study was the number of voting points a team received in a given week. The regression model was performed for each poll (ESPN/USA Today and AP) for each week of the 2003–2004 college football season. Instead of just using the rankings of 1–25 in the polls, the dependent variable became the combined number of votes received by each team, with the top-ranked team receiving 25 points, the second-ranked team receiving 24 points, etc. from each of the voters in the polls. This form of the voting points variable allows for a distinction of the relative difference in quality between two teams in the rankings, proxied by the points, and allows for teams which are not in the top 25 to still be considered, based on the number of voting points they received in a given week.

The independent variables in the model included an intercept, the votes the team received in the poll from the previous week, a variety of television coverage variables (to proxy for games which voters could watch), and the variable of interest, the "Pointspread Differential." The "Pointspread Differential" was defined as the favorite minus underdog score, minus the point-spread on the game (a positive number in terms of the favorite). Positive numbers represented that the favorite team exceeded expectations (won by more than the point-spread) and negative numbers represented a team that failed to meet expectations (won by fewer than the point-spread or lost the game outright).

In terms of the Pointspread Differential variable, the regression results revealed that this variable had a positive and significant effect on the number of votes a team would receive. In both regressions, ESPN/USA Today votes as dependent and AP votes as dependent, this variable was found to be significant at the 1 percent level. In the coaches' poll (ESPN/USA Today), each additional point of the Pointspread Differential was found to increase voting points by slightly more than 1.6 votes. In the writers' poll (AP), similar results were found, as each additional point of the Pointspread Differential was found to increase votes by slightly over 1.75 votes. Overall, the rankings of teams by voters were found to be affected by their performance compared to expectations (the point-spread on the game).

TV coverage was found to play an important role in this study as well, as wins on television tended to bring more votes, and losses on television severely hurt teams in polls. In further regressions within the paper where more detailed TV analysis was considered (dummy variables by network), the Pointspread Differential retained its importance as it was still found to have a positive and significant effect at the 1 percent level with similar magnitude of the coefficients for both the coaches' and writers' polls.

Overall, this study revealed that the point-spread generated in the simple financial market for sports wagering can be useful in explaining aspects of human behavior outside of gambling. This price (point-spread), which is formed in a prediction market, can be used to capture information about expectations. In the case of the college football polls, these expectations are useful in explaining why some teams win and rise in the polls (they win and cover the point-spread, exceeding expectations), other teams win, yet drop in the polls (they win but do not cover the point-spread, failing to meet expectations), and some losses are not as damaging as others (a team loses, but covers the point-spread – they outperform expectations and are rewarded for it).

Similar to prediction markets in the worlds of business, politics, weather, etc., there is potential for the application of data generated by sports gambling markets to improve outcomes by those in indirectly related fields that are influenced by the actual events being forecast. Prices formed in betting markets may help us better understand game attendance, television ratings and other areas of the sports world. As more and better data begins to emerge, including detailed data on betting percentages, volume of bets and even tracking of individual bettor behavior, the likelihood we will learn more about this market and create new applications is very high.

As this field of research continues to evolve and develop, we may observe applications to sports marketing, advertising, merchandise sales and other areas researchers and practitioners are only now beginning to imagine. We believe it is an exciting time to be involved with prediction markets and with the economics and finance of sports. We eagerly look forward to further development and future breakthroughs in this branch of research.

References

Buraimo, B. and Simmons, R. (2009) "A Tale of Two Audiences: Spectators, Television Viewers, and Outcome Uncertainty in Spanish Football." *Journal of Economics and Business*, 61: 326–338.

Carmichael, F., Millington, J. and Simmons, R. (1999) "Elasticity of Demand for Rugby League Attendance and the Impact of 'BSkyB'." *Applied Economics Letters*, 6(12): 797–800.

Dawson, A. and Downward, P. (2005) "Measuring Short-Run Uncertainty of Outcome in Sports Leagues: A Comment." *Journal of Sports Economics*, 6: 303–313.

Forrest, David and Simmons, Robert (2002) "Outcome Uncertainty and Attendance Demand in Sport: The Case of English Soccer." *The Statistician*, 51(2): 229–241.

Forrest, David, Simmons, Robert and Buraimo, Babatunde (2005) "Outcome Uncertainty and the Couch Potato Audience." *Scottish Journal of Political Economy*, 52(4): 641–661.

Knowles, Glenn, Sherony, Keith and Haupert, Mike (1992) "The Demand for Major League Baseball: A Test of the Uncertainty of Outcome Hypothesis." *The American Economist*, 36(2): 72–80.

Lemke, Robert J., Leonard, M. and Tlhokwane, Kelebogile (2010) "Estimating Attendance at Major League Baseball Games for the 2007 Season." *Journal of Sports Economics*, 11: 316–323.

Morley, B. and Thomas, D. (2007). "Attendance Demand and Core Support: Evidence from Limited-overs Cricket." *Applied Economics*, 39: 2085–2097.

Owen, D. and Weatherston, C. (2004) "Uncertainty of Outcome and Super 12 Rugby Union Attendance: Application of a General-to-Specific Modeling Strategy." *Journal of Sports Economics*, 5: 347–370.

Paul, Rodney J. and Weinbach, Andrew P. (2007) "The Uncertainty of Outcome and Scoring Effects on Nielsen Ratings for Monday Night Football." *Journal of Economics and Business*, 59(3): 199–211.

Paul, Rodney J., Weinbach, Andrew P. and Coate, Patrick (2007) "Expectations and Voting in the NCAA Football Polls: The Wisdom of Point Spread Markets." *Journal of Sports Economics*, 8(4): 412–424.

Paul, Rodney J., Weinbach, Andrew P., Borghesi, Rick and Wilson, Mark (2009) "Using Betting Market Odds to Measure the Uncertainty of Outcome in Major League Baseball." *International Journal of Sport Finance*, 4(4): 255–263.

Paul, Rodney J., Weinbach, Andrew P. and Wilson, M. (2010) "The Uncertainty of Outcome in NCAA Football in the Age of the BCS." Working paper, currently under submission.

Peel, David and Thomas, Dennis (1992). "The Demand for Football: Some Evidence on Outcome Uncertainty." *Empirical Economics*, 17(2): 323–331.

Peel, David and Thomas, Dennis (1996) "Attendance Demand: An Investigation of Repeat Fixtures." *Applied Economics Letters*, 3: 391–394.

Rascher, D. (1999) "The Optimal Distribution of Talent in Major League Baseball." In L. Hadley, E. Gustafson and J. Fizel (eds.), *Sports Economics: Current Research* (pp. 27–54), Westport, CT: Praeger.

Schmidt, Martin and Berri, David (2001) "Competitive Balance and Attendance: The Case of Major League Baseball." *Journal of Sports Economics*, 2(2): 145–167.

Theil, H. (1967) *Economics and Information Theory*, McGraw-Hill, Chicago.

Welki, Andrew M. and Zlatoper, Thomas J. (1999) "US Professional Football Game-Day Attendance." *Atlantic Economic Journal*, 27(3): 285–298.

13 Experimental prediction and pari-mutuel betting markets

Charles Noussair[1]

1 Introduction

In the last few decades, experimental methods have gained acceptance in economics as an empirical research methodology. An experiment refers to an incentivized economic situation created by the researcher to address one or more specific research questions. A study can employ either laboratory or field experimental methods. Laboratory experimentation refers to conformity to a set of accepted procedures, which include carefully controlling the interaction and information available to participants. The participants are typically university students, who interact in the physical setting of a classroom or a computer room. Field experimental methods relax some of the constraints on procedures, such as using non-student subjects, allowing larger monetary payments, using a venue other than the laboratory, or framing the decision situation in a less artificial context.

Experimental tests offer an opportunity to investigate some specific aspects of prediction markets that would otherwise be difficult to study. Experimental methods allow the researcher to construct a small-scale model of a new market design for evaluation before it goes live. Experimentation allows the researcher to observe variables that are unobservable otherwise, such as the private information traders have about the probability of future outcomes. One or more parameters of the market can be systematically and exogenously varied, while holding the other parameters constant. At low cost, many independent markets with an identical underlying structure can be created. This allows the researcher to conduct hypothesis tests with more power, to consider how variable the outcomes of a market are, or to evaluate the likelihood and consequences of worst-case scenarios. Precise measures of market efficiency and information aggregation can be registered throughout the life of an experimental prediction market. Thus, experiments are well-suited to studying the properties of these markets and how their design might be improved.

In this chapter I review some of the more important experiments that have contributed to the scientific understanding of prediction markets. The emphasis here is not on new applications of prediction markets, but rather on those experiments that focus on understanding how prediction markets operate generally, under what conditions they perform well and how their design can be improved. Overall, the experiments demonstrate the power of prediction markets to aggregate information and make accurate predictions of future events in the form of prices. However, they also illustrate a variety of pitfalls that can arise in such markets. Fortunately, experimental economics also provides a methodology to evaluate potentially better market designs. The chapter is organized in the following manner. Section 2 considers the literature at the intellectual origin of the methods used in experimental research on prediction markets. Section 3 reviews work focused on open contingent-claims prediction markets. Section 4 considers pari-mutuel betting markets, which are closely related. Section 5 offers a brief conclusion.

2 Background: early experiments on market-clearing, long-lived assets and information aggregation

The methodology used in the experimental study of prediction markets has its origin in the development of laboratory techniques to test the theory of perfect competition in markets for "perishable goods." These are characterized by classical demand and supply functions, production to demand and immediate consumption of the good traded. Chamberlain (1948) and Smith (1962) search for conditions under which prices and quantities traded correspond to the competitive equilibrium levels. Smith (1962) employs continuous double-auction trading rules in the markets that he studies. This type of market operates in continuous time; any trader at any time can submit a public price quote to buy or sell one or more units, and an offer can be accepted by any other agent at any time. Smith's work demonstrates that if the market is organized properly, it is possible to make prices and quantities in a single market converge reliably to the competitive equilibrium levels, and to achieve efficient outcomes. The result has been generalized to more complicated environments and to various different market rules, indicating that the conditions required for achieving competitive equilibria in perishable goods markets are not very constraining. The market reveals the competitive equilibrium price, which allows efficient coordination of exchange, and individuals can be induced to behave like price takers, even if there are only a small number of traders. Continuous double-auction rules have become the standard institution of exchange used in the experimental study of market behavior.

Some of this research focuses on whether analogous properties of markets exist in markets for assets. While no single property always distinguishes a perishable good from an asset, one of the following distinctions is typically made in experimental research:

1. Assets are durable and have a life of multiple periods, while perishable goods have no value beyond the current period.
2. Asset markets allow for speculative behavior; the same individual can purchase and sell, and whether an agent is a buyer or a seller depends on

prevailing market conditions at the moment. Thus, traders in asset markets have limit prices that evolve as agents' expectations change over time, while limit prices are typically fixed in goods markets.

3 In asset markets, there are often insiders who have better information about the value of the asset than other agents have, while goods typically have a known private value or cost to each agent.

The earliest experimental studies of multi-period assets are those of Forsythe *et al.* (1982, 1984) and Friedman *et al.* (1984). These studies describe the behavior of experimental markets for assets with a life of two and three periods (a period is defined as a unit of time between two dividend payments). Though these horizons are short, traders do face a situation where they have incentives to arbitrage intertemporally and to form expectations about prices in future periods. After a number of replications of two- or three-period asset markets, prices in the last period converge to approximately the rational expectation equilibrium level. However, convergence is slower and less reliable for period prices, the longer the period precedes the final one. Rational expectation equilibrium prices cannot be discovered until the price for the last period stabilizes, and the price discovery process unravels backward. The presence of futures markets aids and accelerates convergence to rational expectations equilibrium (Forsythe *et al.*, 1984; Friedman *et al.*, 1984).

However, markets for longer-lived assets have a strong tendency to generate price "bubbles." This result is originally due to Smith *et al.* (1988), but it has been widely replicated and shown to be robust to numerous modifications of the experimental design. In the original design of Smith *et al.*, markets are created for assets with a life of a finite number of periods (usually 15 or 30 periods). The asset pays a dividend in each period, which (other than in a few sessions where there is a final fixed terminal value for the asset) is the only source of intrinsic value. The dividend payment is identical for all traders and the distribution of dividends is common knowledge to all traders. The time-series of transaction prices in markets with this structure does not track the fundamental value, but rather is characterized by price bubbles and crashes. A bubble is an extended period of time, during which prices are much higher than fundamental values, while a market crash is a sudden and rapid fall in prices.

Some factors are known to mitigate bubbles in experimental markets. Haruvy and Noussair (2006) have shown that allowing short-selling reduces prices, but if the short-sale constraints are too loose, prices are below fundamental values. They also replicate an earlier result by Caginalp and Smith (1998), who show that the more liquidity traders have available to make purchases, given the fundamental value and the total stock of the asset, the higher the price level relative to fundamentals. If short-sale constraints are set at an appropriate level, which may or may not be plausible given the information available to a market designer, the possibility of making short-sales can push prices toward fundamentals (Ackert *et al.*, 2006). Noussair and Tucker (2006) show that the addition of a sufficient number of derivative futures markets has a strong tendency to reduce, and indeed often completely eliminate, spot market bubbles and crashes. Haruvy *et al.* (2007) show that when individuals are required to predict (privately) the prices that will occur in the future, individuals do not anticipate bubble and crashes, but rather tend to extrapolate previous trends. Thus, crashes are typically a surprise to market participants. However, all of the studies that examine changes in asset market behavior as traders acquire more experience find that market prices are closer to fundamentals, the more experience traders have in the same environment.

In addition to the fact that they typically trade long-lived assets, prediction markets have the feature that some traders have better information about the value of the asset than others. One function of a market, and indeed the primary purpose of a prediction market, is to reveal this information with the market price. Several early experimental studies indicate that markets have a strong tendency to disseminate private information, provided that enough individuals hold the information. For assets that have a life of only one period, and have a common though uncertain value, Plott and Sunder (1982) observe that when insiders who know the true value of the asset are present, prices in continuous double-auctions reveal the insider information. This result shows that there exist conditions where it is possible to use a decentralized market to disseminate privately held information.

Later studies illustrate the limits of the ability of markets to reveal information. Plott and Sunder (1988) study the issue of whether markets can aggregate privately held information. They endow insiders with a portion of – but not all of – the information needed to determine the true value of the asset. Only the aggregation of all of the information held by insiders would allow the state of nature, and therefore the fundamental value of the asset, to be deduced with certainty. The results on information aggregation are mixed. In a setting in which there are markets for contingent claims and in which dividends differ between agents, prices tend toward the level corresponding to rational expectations. However, when only one security is exchanged, prices do not correctly reflect the available information. Forsythe and Lundholm (1990) show that, for the same environment, sufficient trader experience, in conjunction with common knowledge of payoffs, enables the market to reliably aggregate and reveal the inside information.

Another observation emerging from early experimental research is that markets may price as if they reveal information that is not actually held by any traders. If the presence of insiders is uncertain, market activity can lead to convergence of prices to levels that are consistent with the presence of insiders, even when no insiders actually exist. This occasional failure of markets to reveal the absence of information is termed an "information mirage" (Camerer and Weigelt, 1991). Individuals may trade on the basis of inferences they make from the trades of others, creating price paths that falsely reveal information that traders do not actually have. Camerer and Weigelt usefully distinguish between mirages, which are caused by uncertainty about the information of others, and bubbles, which appear to be caused by uncertainty about the rationality of others.

The experimental literature on prediction markets can be classified as belonging to one of two branches. One branch, discussed in Section 3, focuses on the operation of open contingent-claims markets, in which agents buy and sell securities that payoff in the event that specific future events occur. The prices are taken as measures of the likelihood of events or the expectation of the magnitude of an outcome variable to be realized in the future. The second branch of the literature, which is the topic of Section 4, discusses research on pari-mutuel betting markets. These are markets in which traders can make irrevocable bets on a future event, and the market odds change in response to betting behavior. The odds are interpreted as a measure of the likelihood of future events. Contingent-claims and pari-mutuel markets have a close theoretical link. However, because of differences in framing, in the institutions of price formation typically present, and in the revocability of purchases, they are discussed separately here.

3 Contingent-claims prediction markets

We begin by describing the structure of a standard experimental prediction market. Consider the following. There are two possible outcomes, *A* and *B*, and the outcome is determined and publicly announced at time *t*. A market is open, before time *t*, to trade securities which earn a liquidation value at time *t*. The liquidation value is based on whether the outcome is *A* or *B*. In the most common version, one security, which we will refer to as security *A*, pays out 1 at time *t* if the outcome is *A*, and 0 if the outcome is *B*. Analogously, a security *B* is created, which pays 1 at time *t* if the outcome is *B*, and 0 if the outcome is *A*. Markets are created to exchange each of these securities, and typically the markets are operating simultaneously and using continuous double-auction trading rules. The price of security *A* at any time *t* is interpreted as the probability "the market" assesses that the outcome will be *A*, given the information market participants hold at time *t*. In principle, any finite number of securities can be specified, and the price of each can be viewed as the probability of a separate event. If the list of possible events spanned by the securities is not exhaustive, an additional security, paying out 1 in the event that none of the other outcomes occurs, and 0 otherwise, can be specified. Adding this "residual" asset has the advantage that it induces a transparent no-arbitrage condition: that the sum of the prices of all of the securities equals 1. If the sum of the prices is <1 (respectively >1), an arbitrageur can purchase (resp. sell) the market portfolio and make a riskless profit. The market price of such contingent claims has been taken by many researchers and practitioners to correspond to the average belief of the traders in the market.[2]

The structure of the set of contingent claims described above has what is referred to as an "all-or-nothing" feature. However, many interesting applications have the property that they have a continuum of outcomes, such as the percentage of votes a political party receives in an election, a stock price or a macroeconomic variable. In such cases, a proportional share market can sometimes be designed. A security A' can be created, which pays out in proportion to the value of the outcome variable. For example, in an election, A' can pay 0.01 at time *t* for each percentage point of the vote that candidate *A* receives, and similarly for another security B' representing the share for an opposing party. A final security "Not A' or B'" can capture the vote share that goes to parties other than A and B. In the case of predicting a stock price, a menu of all-or-nothing claims can be defined with the following structure. One security can payoff 1 if the stock price at future date t lies between p_1 and p_2, and 0 otherwise; another can payoff 1 if and only if the price at t is between p_2 and p_3, etc.[3]

Perhaps the most celebrated series of contingent-claims prediction market experiments is the Iowa Political Stock Market (Forsythe *et al.*, 1992), a field experimental platform used to predict winners and vote-shares in elections. The experiments consist of either all-or-nothing markets, proportional share markets or both types operating simultaneously. The 1988 US presidential election contest, between the two major party candidates, George H.W. Bush and Michael Dukakis, was the market's first major success. Its proportional share markets generated vote-share predictions differing from the final vote shares by less than 0.1 percent for Bush and 0.2 percent for Dukakis. These were closer to the final percentage than the final polls of any of the major polling agencies. The success was repeated in 1992 for a three-way presidential race between William Clinton, George H.W. Bush and Ross Perot. The election-eve forecast implied by the market prices for each of the three candidates differed from the actual fractions received by an average absolute error of 0.2 percent, again much closer than the election-eve polls. The success of the Iowa Political Stock Market has stimulated a line of laboratory research focusing on the reasons for its accuracy. This research has isolated several key factors influencing the level of accuracy of prediction markets.

3.1 Accurate pricing relies on marginal traders

Forsythe *et al.* (1992) conjecture that the efficient operation of the market relies on the presence of *marginal traders*. While the average trader might be subject to a variety of decision biases, these marginal traders can assure accurate pricing if they have sufficient weight in the market to be decisive. A representative sample of the population is not critical for the operation of a political stock market, as long as enough marginal traders with enough resources are participating. Marginal traders are characterized by the following properties. First, they are market makers instead of price takers. Second, they take advantage of other traders' biases to earn profits. Forsythe *et al.*'s operational definition for marginal traders are those who (1) submit a limit order at the end of the day at a price within 2¢ of the last traded price of the day; or (2) who make offers that are accepted some time during the day. Using this criterion, Forsythe *et al.* (1992) identify 22 marginal traders out of a total of 192 traders. These traders appear to arbitrage effectively among biased traders, adjust more quickly to relevant new developments in the campaign, invest more money than others and trade more shares than the average trader. Oliven and Rietz (2004) adopt a narrower definition of marginal traders, essentially consisting of those individuals who tend to submit price quotes that later result in trades. They find that, on average, these traders are less prone to forego arbitrage opportunities, and have more education and market-related experience than the rest of the trader population. This suggests that prediction markets have the feature that they induce the more rational traders to set the prices, that this feature enhances a market's ability to price accurately, and that it is a key to the successes of political stock markets.

3.2 Manipulation of prediction markets is difficult

What if there are individuals who are willing to incur losses to distort prices in a prediction market? This is a possibility, for example, in a widely publicized political stock market if the market prices could affect the expectations or preferences of voters (see, for example, Bohm and Sonnegard, 1999). However, the experimental evidence available suggests that such manipulation is very difficult. Camerer (1998) conducts a field experiment at a racetrack pari-mutuel betting market. He studies whether placing large bets and withdrawing them at the last minute affects the betting behavior of other, subsequent, bettors. The effect is statistically insignificant, indicating that manipulation of the market is difficult.

Hanson *et al.* (2006) conduct a laboratory experiment replicating the design of Plott and Sunder (1988), but including some traders who have incentives to manipulate prices to a target level. They find that the manipulators are unable to distort prices. Those traders without manipulation incentives step up and act as counterparties and thereby compensate for the actions of the manipulators. This is consistent with the available empirical evidence from non-experimental betting markets, which also indicates that manipulation is difficult and exerts at most a short-term impact (Rhode and Strumpf, 2004, 2007; Wolfers and Zitzewitz, 2004). Although Hansen *et al.* (2004) document an interesting case of manipulation lasting ten days in a prediction market organized for the 1999 Berlin state elections, such behavior seems to be the exception rather than the rule.

3.3 Bubbles and violation of no arbitrage might occur

The research on asset markets described in Section 2 noted that elevated prices are common in experimental asset markets. Research has shown this to be a legitimate concern in prediction markets. This is best illustrated in the laboratory experiment of Rietz (2005). In his experiment, there are two contracts traded in a winner-takes-all prediction market. One is called the "green certificate," the other the "blue certificate." Subjects are endowed with experimental currency that they can use to purchase the certificates. At the end of the trading period, green certificates are liquidated at 1,000 (units of experimental currency), with an *ex ante* probability of 0.7, whereas blue certificates are liquidated at 1,000, with a probability of 0.3. The certificate that pays out is determined at the end of each period by a random draw of a marble from a bucket containing six blue and 14 green marbles. This probability distribution and corresponding payoff are common knowledge. Therefore, the equilibrium prices are 700 and 300 for blue and green, respectively. An arbitrage opportunity exists whenever the sum of the two contract prices was strictly greater or strictly less than 1,000.

There are several treatment conditions. The baseline treatment is called OPI, in which simultaneous continuous double-auction markets operate for the two certificates. Additional treatments are conducted with the goal of facilitating arbitrage through providing additional training for subjects. In the OPI(M) treatment, the experimental instructions are modified to stress that the unit portfolio (consisting of one unit of each certificate) is always worth 1,000 independently of the marble drawn. Additional instruction is provided part-way through the sessions, which teach traders to identify arbitrage opportunities. In the OPI(E) treatment, all traders have prior experience in similar markets and this fact is common knowledge. In OPI(R), the periodic feedback of which marble is drawn in each period is not provided, in contrast to the other treatments. In all of these treatments, the securities are systematically overpriced, even in the last few periods. The sum of the two prices averages 1,242, an overpricing of 24.2 percent, relative to the no-arbitrage restriction. Individual security prices exceed theoretical benchmarks in 90 percent of periods.

In the OPI(P) treatment, direct trading of the unit portfolio is made possible by opening up an extra market to trade it. This reduces the average overpricing to 13.7 percent, and the percentage of market periods with overpricing to 76 percent of periods. In the OPI(S) treatment, short-selling is permitted, and this also reduces prices modestly, though they remain overvalued by 18.4 percent on average and exceed the theoretical level in 75 percent of periods. In the last treatment, OPI(A), the experimenter plays the role of an active arbitrageur who sells the unit portfolio if the sum of bids is larger than 1,000 and buys the unit portfolio if the sum of asks is below 1,000. This treatment does significantly reduce the level of aggregate overpricing. However, the arbitrageur does not improve the efficiency of individual security prices. Rather, his activity pushes the ratio of the prices of the two certificates away from theoretical levels, increases trading volume and increases price volatility. To bring the prices down, the arbitrager must be very active, participating in 60 percent of all trades.

Nevertheless, in all but the OPI(A) treatments, relative prices, the ratio of the prices of the two certificates, are close to the theoretical prediction. Thus, relative prices provide an accurate prediction of outcome probabilities after prices are appropriately normalized. Rietz (2005) concludes that normalized (relative) prices "are tied more closely to state probabilities than absolute prices." He suggests that the degree of the systematic overpricing could be explained by a bias toward purchases and away from sales.

3.4 False consensus and wishful thinking can distort outcomes

In principle, traders' personal preferences over outcomes should not influence market behavior. However, there is evidence that wish-fulfillment exerts an influence on outcomes. Analogously to the tendency of sports bettors to bet on home teams, there is a tendency for individuals trading in political stock markets to hold shares in the candidates they favor. This effect appears to have two aspects. The first is a *false consensus effect*. A trader overestimates the chance of his preferred candidate or political party winning, because he overestimates the proportion of voters that agree with him. The other is the *assimilation-contrast effect*. This is a tendency for a trader to interpret new information in a manner that reflects positively on his preferred candidate or party.

Evidence for the assimilation-contrast effect is provided by Forsythe *et al.* (1999). During the 1988 election campaign, Bush supporters increased their holdings of the security that corresponded to Bush following each of the three televised debates between the two candidates, while Dukakis supporters were net sellers of Bush after each debate. In a survey, 96.4 percent of the traders who preferred Bush responded that Bush performed at least as well as his competitors did in the last debate. Meanwhile, 80.4 percent of Dukakis supporters thought Dukakis was at least as good as his competitors. Over the course of the campaign, Bush supporters increased their holdings of Bush shares by an average of 1.12 units and decreased their holding of Dukakis shares by 2.26 units. The opposite pattern held for Dukakis supporters.

Forsythe *et al.* (1999) also conduct a laboratory experiment, in which traders are endowed with preferences over the outcomes. The experiment design is the same as the baseline treatment in the Rietz (2005) study, with the exception that traders can receive a bonus of 4,000, depending on the final state. In some periods, there is no bonus awarded; in other periods the prize is given to all traders if the "blue" state occurs; and in the remaining periods all traders receive the bonus if the "green" state is realized. The prize a trader receives does not depend on how many certificates he holds or on the price and volume of trade of each certificate. Thus, the theoretical prediction is that there are no differences in prices across the three types of period. However, if the wish fulfillment effect is strong, the market would overprice the certificate corresponding to the state currently yielding a bonus. Indeed, prices do increase for the security that corresponds to the state yielding the bonus in the current period, demonstrating that wish fulfillment can influence market outcomes.

4 Pari-mutuel betting markets

A number of interesting experiments have explored the behavior of pari-mutuel betting markets. These are markets in which individuals have an opportunity to place bets on one or more outcomes, and the bets determine the betting odds. Once bets are placed, they cannot be revoked. Most of the studies focus on settings in which individuals have private information correlated with the eventual outcome, and on the issue of whether this private information is aggregated into the odds. At the level of individual decision-making, a primary focus is to study the conditions under which players will engage in *herding* and in *contrarian* behavior.

Herding is defined as betting in disagreement with ones private signal but in favor of the consensus based on prior bets. Herding is, essentially, "going with the crowd" and against one's own information. This can be optimal to do if the informational content of the prior bets weighs in favor of one outcome to a greater degree than one's private signal weighs in the favor of another outcome. However, it can also be incorrect to herd, if the information implicit in others' bets does not outweigh the information contained in the private signal for a rational agent. Contrarian behavior involves betting against one's own private information and against the consensus. Contrarian behavior is of special interest, because it can generate or accentuate the favorite–longshot bias. This is a tendency for betting odds to overstate the probability of a longshot being the outcome, making it more profitable in expectation for other bettors to bet on a favorite than on a longshot. This is because the market odds reflect contrarian bets and overprice longshots, as well as underpricing favorites. Along with contrarian behavior, various plausible betting heuristics, such as betting with equal probability on each outcome regardless of payoff and perceived winning probability, or betting based on idiosyncratic tastes for the outcome (favorite number, preferred color of horse, acquaintance with jockey, etc.) can also accentuate the favorite–longshot bias. A favorite–longshot bias can also arise as a consequence of the transformation of probabilities (Kahneman and Tversky, 1979; Prelec, 1998), risk-seeking preferences, or the presence of utility for beating the odds by betting on a successful longshot.

The source of the favorite–longshot bias has been the specific topic of a number of experimental studies. Piron and Smith (1995) report an experiment that they interpret as supporting the idea that the transformation of probabilities, rather than the existence of utility from beating the odds, is a cause of the favorite–longshot bias. Hurley and McDonough (1995) consider whether the favorite–longshot bias is a result of the fact that the racetrack earns a fraction of the amount bet. They present a model in which, with zero take for the market maker, there is no bias, but a positive take generates a bias. However, their experiment, which compares markets with and without market maker costs, yielded no difference between the two treatments, and thus their model was not supported.

Drehmann *et al.* (2005) conduct a large internet-based field experiment, in which players play a betting game with the structure of a pari-mutuel betting market. Players endowed with private information are offered, sequentially, one opportunity to place a bet on one of two possible outcomes, or to refrain from betting. The odds, the prices for a bet on each outcome, are updated after each bet so that they reflect the conditional probability of each outcome. Thus, the price equates the expected payoff of betting on each of the two available alternatives based on the public information only. Therefore, it is always optimal to bet in agreement with one's private information. There are a number of treatments that vary the displays that bettors are presented with, whether an option not to bet is available, and whether prices are set assuming the presence of error in the bets of prior bettors. However, all of the treatments have in common the feature that it is always optimal, in terms of maximizing expected value, to bet on the outcome that is in agreement with one's private signal. Despite this, only two-thirds of decisions are consistent with private signals. There is little herding, but abstention from making a bet, as well as contrarian behavior, is common.

Similar results are obtained by Cipriani and Guarino (2005). They also report an experiment, in which each bettor in a sequence possesses some private information about whether an asset's value is likely to be 0 or 100. They can bet on the outcome by either buying or selling based on the information they have and (in some treatments) on the history of betting decisions of prior bettors in the sequence. They conduct one treatment in which the price is fixed at 50, and one in which it is flexible, varying according to prior betting activity in a manner analogous to Drehmann *et al.*'s (2005) experiment. In the flexible price treatment, it is always optimal to bet in favor of one's private signal. Indeed, they find that the flexible price condition leads to a lower incidence of herding than the fixed price treatment. However, they observe a high incidence of abstention from betting and of contrarian behavior. The results are similar whether or not the history of preceding bets is provided. In both this study and in Drehmann *et al.* (2005), contrarian betting impedes the ability of the market to aggregate information.

Koessler *et al.* (2010) introduce a design feature that greatly reduces the incidence of contrarian behavior. They construct a pari-mutuel market in which players move in a fixed sequence, as in the last two studies. However, at the time each individual makes a bet, all players must submit a belief assessment about the state. That is, they much assess the probability that each of the two outcomes will be realized. Beliefs are remunerated based on how far they are from the actual outcome, according to a function that ensures that it maximizes expected payoff for an agent to truthfully report his actual belief. Koessler *et al.*'s experiment has three treatments. In the Bet treatment, individuals only submit bets. In the ObsPred treatment, one group of players submits bets, and another group of observers, endowed with private signals and who can observe the history of trades, submits beliefs. In the BetPred treatment, the same players make bets and submit beliefs. When bettors submit beliefs, in BetPred, contrarian behavior decreases sharply compared to the other treatments. Under BetPred, the market also aggregates more information, primarily by reducing contrarian betting, and exhibits a much smaller favorite–longshot bias. Beliefs are also more accurate when bettors, rather than observers, are submitting them. It appears that eliciting beliefs from bettors directs more of their attention to the probability of each outcome eventually being realized. This may cause relatively less weight to be placed on the high payoff associated with the longshot in the (unlikely) event that it wins, reducing the tendency to bet on it.

The pari-mutuel market experiments described above are highly structured in terms of the precise sequencing of bets, and the constraint that each individual can only submit one bet. Such structure facilitates the study of individual decisions, and the testing of theoretical models of decision-making. In many real pari-mutuel betting markets, however, such specific structure is not present. Rather, the market is open-ended, and an individual can place many bets at the timing of her choosing. The laboratory experimental study of open continuous pari-mutuel betting markets was initiated by Plott *et al.* (2003). In their experiment, they conduct several markets that operate simultaneously. In each market, the experimenter sells contingent claims on a different outcome. Individuals are endowed with some private information about the likelihood of the outcomes and a fixed budget with which they can purchase tickets. While the market is open, any individual can, at any time, purchase as many tickets as he wishes to at a fixed price per ticket. All ticket purchases are public information, each market shows the number of tickets it has left for sale and the odds are posted periodically. Tickets purchased cannot be resold. There are two treatment conditions, "Not Sets," in which individuals' private information would allow them to eliminate some outcomes with certainty, and "PIC," where private information allows updating of probabilities, but not the elimination of any of the possible outcomes. A favorite–longshot bias appears in both settings, though information aggregation is better in Not Sets than in PIC. Strategic behavior, in the form of waiting until late in the period (including the submission of bets just before the market closes) and attempts to bluff and mislead early in the market period, are common and appear to accentuate the favorite–longshot bias.

Axelrod *et al.* (2009) modify the Plott *et al.* (2003) design in two ways. They impose a cost of delay, to encourage earlier betting, with the goal of reducing early-period strategic waiting and bluffing. This is done by increasing the price of each bet at a constant rate over the course of the market period. They also divide the period into two rounds of betting, and after the first round, the current interim odds are posted. The authors find that the process of information aggregation is more rapid with these modifications. The favorite–longshot bias is present in the first round, but largely disappears in the second round, suggesting that it is a disequilibrium phenomenon; that is, a transitory pattern that fades away if the process of market clearing is permitted to continue unimpeded.

Roust and Plott (2005) propose and test a further enhancement of this system. In their betting markets, there are two stages. In the first stage players can buy contingent claims on each outcome with fixed budgets of fiat money (which has no value other than as a means to purchase the claims). Prices are constant over the course of this stage. Because the money has no other use, there is an incentive for individuals to spend all of their budgets. The number of claims purchased is not disclosed until the stage ends, so there is no incentive to bluff and mislead within the first stage. The second stage of the market is a pari-mutuel betting market with regularly increasing prices over time, as in Axelrod *et al.* (2009). This two-stage system reduces the incidence of bubbles and information mirages relative to, and achieves better values of measures of information aggregation than, the systems studied previously by Plott *et al.* (2003) and Axelrod *et al.* (2009).

5 Conclusion

This chapter has reviewed some of the more important experimental studies of prediction markets. The focus is on issues of market performance in terms of generating good predictions that reflect the private information potential participants hold. The data give an encouraging, though qualified, picture of prediction market performance. A blanket claim that such markets will always perform well is not supported. However, the creation of a market will typically aggregate some privately held information if such information is available, and thus a prediction market is typically beneficial. The accuracy of a prediction market appears to depend considerably on how demanding the information aggregation task is, on how active and sophisticated traders are, and on features of the way the market trading process is organized.

Some results from experimental research appear to be rather robust and may eventually come to be recognized as general principles. Market prices appear to have a tendency to stabilize, and in sufficiently simple and stationary environments, this stabilization tends to occur at close to a fundamental value. However, the presence of marginal traders with sufficient weight in the market is critical to good performance. Traders with repeated experience under the same conditions create conditions more conducive to accurate pricing. Arbitrage cannot necessarily be counted upon to lead to efficient pricing. Market price levels are sensitive to non-fundamental factors: greater liquidity tends to raise prices and loose short-sale constraints lower prices. Therefore, market bubbles are a distinct possibility, and are more likely as the life of the asset gets longer. Futures markets improve the efficiency of spot markets.

A prediction market can be expected to function better, the less demanding its task. If the private information content is more widely held or more precise, in the sense that some participants can completely eliminate some contingencies, the market has a better chance of integrating the information into the market price. Intentional manipulation of prices appears to be very difficult. However, if the biases of wishful thinking, false consensus and assimilation contrast, are widely held, they can distort prices. Furthermore, while the market can sometimes fail to reveal some privately held information, it can also appear to the observer to reveal fictitious information that is not actually there.

The research on pari-mutuel betting confirms that the favorite–longshot bias is a general phenomenon that appears even in simple markets. It appears to have a multiplicity of causes: a bias toward contrarian betting; a failure to bet even when one has useful information; a bias toward placing equal weight on all betting options; strategic bluffing to influence future odds; and waiting to bet to avoid revealing what one knows. Changes in the rules of the betting market can reduce the favorite–longshot bias. Thus, there is scope for good institutional design to improve market outcomes, and for poor designs to magnify inefficiencies.

Notes

1 I am indebted to Yi Long Xu for his very competent research assistance on this project.

2 Manski (2006) discusses the hazards of interpreting the price as the belief of an average agent, and shows that the price instead defines a bound on the average belief. Wolfers and Zitzewitz (2007) have shown that market prices may provide somewhat biased estimates of average beliefs, but that the biases are generally not severe.

3 The choice of the menu of securities is not innocuous for the inference of outcome probabilities from market prices. Sonneman *et al.* (2008) suggest that there is a natural bias toward placing equal weight on all alternatives (a $1/N$ bias). They report several

experiments in which the probabilities of events, implicit in the market prices in contingent-claims markets, depend on the way the outcome space is partitioned into different events. The price of a security payoff in the event of x plus the price of a security payoff under y is greater than one that pays out in the event of either x or y.

References

Ackert, L.F., N. Charupat, B.K. Church and R. Deaves (2006), "Margin, Short Selling, and Lotteries in Experimental Asset Markets," *Southern Economic Journal*, 73(2), 419–436.

Axelrod, B.S., B.J. Kulick, C.R. Plott and K.A. Roust (2009), "The Design of Improved Pari-mutuel-type Information Aggregation Mechanisms: Inaccuracies and the Long-shot Bias as Disequilibrium Phenomena," *Journal of Economic Behavior and Organization*, 69(2), 170–181.

Bohm, P. and J. Sonnegard (1999), "Political Stock Markets and Unreliable Polls," *Scandinavian Journal of Economics* 101(2), 205–222.

Caginalp, G., D. Porter and V. Smith (1998), "Initial Cash/Asset Ratio and Asset Prices: An Experimental Study," *Proceedings of the National Academy of Sciences* 95(2), 756–761.

Camerer, C. (1998), "Can Asset Markets Be Manipulated? A Field Experiment with Racetrack Betting," *Journal of Political Economy*, 106(3), 457–482.

Camerer, C. and K. Weigelt (1991), "Information Mirages in Experimental Asset Markets," *Journal of Business*, 64, 463–493.

Chamberlain, E.H. (1948), "An Experimental Imperfect Market," *Journal of Political Economy*, 56(2), 95–108.

Cipriani, M. and A. Guarino (2005), "Herd Behavior in a Laboratory Financial Market," *American Economic Review*, 95(5), 1427–1443.

Drehmann, M., J. Oechssler and A. Roider (2005), "Herding and Contrarian Behavior in Financial Markets: An Internet Experiment," *American Economic Review*, 95(5), 1403–1426.

Forsythe, R. and R. Lundholm (1990), "Information Aggregation in an Experimental Market," *Econometrica*, 58(2), 309–340.

Forsythe, R., T. Palfrey and C. Plott (1982), "Asset Valuation in an Experimental Market," *Econometrica* 50(3), 537–568.

Forsythe, R., T. Palfrey and C. Plott (1984), "Futures Markets and Informational Efficiency: A Laboratory Examination," *Journal of Finance*, 39, 955–981.

Forsythe, R., F. Nelson, G.R. Neumann and J. Wright (1992), "Anatomy of an Experimental Political Stock Market," *American Economic Review*, 82(5), 1142–1161.

Forsythe, R., T. Rietz and T. Ross (1999), "Wishes, Expectations and Actions: A Survey on Price Formation in Election Stock Markets," *Journal of Economic Behavior and Organization*, 39, 83–110.

Friedman, D., G. Harrison and J. Salmon (1984), "The Informational Efficiency of Experimental Asset Markets," *Journal of Political Economy*, 92(3), 349–408.

Hansen, J., C. Schmidt and M. Strobel (2004), "Manipulation in Political Stock Markets: Preconditions and Evidence," *Applied Economics Letters*, 11, 459–463.

Hanson, R., R. Oprea and D. Porter (2006), "Information Aggregation and Manipulation in an Experimental Market," *Journal of Economic Behavior and Organization*, 60(4), 449–459.

Haruvy, E. and C.N. Noussair (2006), "The Effect of Short-Selling on Bubbles and Crashes in Experimental Spot Asset Markets," *Journal of Finance*, 61(3), 1119–1157.

Haruvy, E., Y. Lahav and C.N. Noussair (2007), "Traders' Expectations in Asset Markets: Experimental Evidence," *American Economic Review*, 97(5), 1901–1920.

Hurley, W. and L. McDonough (1995), "A Note on the Hayek Hypothesis and the Favorite Longshot Bias in Parimutuel Betting," *American Economic Review*, 85(4), 949–955.

Kahneman, D. and A. Tversky (1979) "Prospect Theory: An Analysis of Decision under Risk," *Econometrica*, 67, 263–291.

Koessler, F., C. Noussair and A. Ziegelmeyer (2010), "Information Aggregation and Belief Elicitation in Experimental Parimutuel Betting Markets," working paper, Tilburg University.

Manski, C. (2006), "Interpreting the Predictions of Prediction Markets," *Economics Letters*, 91, 425–429.

Noussair, C. and Tucker, S. (2006), "Futures Markets and Bubble Formation in Experimental Asset Markets," *Pacific Economic Review*, 11(2), 167–184.

Oliven, K. and T. Rietz (2004), "Suckers are Born but Markets are Made: Individual Rationality, Arbitrage, and Market Efficiency on an Electronic Futures Market," *Management Science*, 50(3), 336–351.

Piron, R. and L.R. Smith (1995), "Testing Risklove in an Experimental Racetrack," *Journal of Economic Behavior and Organization*, 27, 465–474.

Plott, C.R. and S. Sunder (1982), "Efficiency of Experimental Security Markets with Insider Information: An Application of Rational Expectations Models," *Journal of Political Economy*, 90(4), 663–698.

Plott, C.R. and S. Sunder (1988), "Rational Expectations and the Aggregation of Diverse Information in Laboratory Security Markets, with Shyam Sunder," *Econometrica*, 56, 1085–1118.

Plott, C.R., J. Wit and W.C. Yang (2003), "Parimutuel Betting Markets as Information Aggregation Devices: Experimental Results," *Economic Theory*, 22, 311–351.

Prelec, D. (1998), "The Probability Weighting Function," *Econometrica*, 66, 497–527.

Rhode, P. and K. Strumpf (2004), "Historical Predictions Markets: Wagering on Presidential Elections," *Journal of Economic Perspectives*, 18(2), 127–142.

Rhode, P. and K. Strumpf (2007), "Manipulating Political Stock Markets: A Field Experiment and a Century of Observational Data," working paper.

Rietz, T. (2005) "Behavioral Mispricing and Arbitrage in Experimental Asset Markets," working paper, University of Iowa.

Roust, K. and C.R. Plott (2005), "The Design and Testing of Information Aggregation Mechanisms: A Two-stage Pari-mutuel IAM," working paper, California Institute of Technology.

Smith, V.L. (1962), "An Experimental Study of Competitive Market Behavior," *Journal of Political Economy*, 70(2), 111–137.

Smith, V.L., G. Suchanek and A. Williams (1988), "Bubbles, Crashes, and Endogenous Expectations in Experimental Spot Asset Markets," *Econometrica*, 56, 1119–1151.

Sonneman, U., C. Camerer, C.R. Fox and T. Langer (2008), "Partition-Dependent Framing Effects in Lab and Field Prediction Markets," working paper, UCLA.

Wolfers, J. and E. Zitzewitz (2004), "Prediction Markets," *Journal of Economic Perspectives*, 18(2), 107–126.

Wolfers, J. and E. Zitzewitz (2007), "Interpreting Prediction Market Prices as Probabilities," working paper, University of Pennsylvania.

14 The economic analysis of sports betting by expert gamblers and insiders

A survey

John Peirson

1 Introduction

Gambling on sporting events has attracted much academic attention over the last 30 years – for reviews, see Thaler and Ziemba (1988), Sauer (1998), Vaughan Williams (1999), Coleman (2004) and Clotfelter (2005). In particular, the impact of the use of information by gamblers has been examined in a large number of theoretical and empirical studies. The purpose of this survey is to consider the literature on gambling by experts and those with insider information. We show that most models in the literature describe equally well the behaviour of gamblers with access to privileged information and those with expert skills. We investigate possible differences in the behaviour of the two types of informed gamblers. These differences are important, as expert gambling is usually regarded as a legal activity, but insider gambling is often considered to be wrong and, in many countries, is against the law.

The survey in five further sections. The second section considers definitions of gambling by experts without access to privileged information and betting by those with inside information. The third section reviews the literature on betting by insiders and experts on sporting events. The fourth section considers whether there are observable differences between the two types of betting. The fifth section investigates the impact of the two types of betting in the context of concerns about the integrity of gambling markets and the internet providing new and different markets for sports betting. The final section gives a conclusion focusing on the need for new theoretical and empirical developments as models of betting by insiders and experts are very similar and there is a generally perceived need for regulation of insider gambling, but not gambling by experts.

2 Definitions of expert and insider gambling

The literature on sports betting has not carefully defined the difference between gambling by experts and those with inside information (Peirson and Smith, 2010). Thus, to consider the definitions of these two forms of gambling, one must turn to the law and economic literature on insider trading. The lawyers Manne (1966) and Bainbridge (2001, p. 3) define insider trading as 'trading …

while in possession of material nonpublic information'. By comparison, 'informed traders [without inside information] include market professionals, analysts, broker-dealers, market makers and sophisticated investors' (Beny, 2002, p. 8). This distinction is useful for the purpose of analysing informed gambling.

This survey suggests that there is an important distinction to be made between bettors who process publicly available information to form accurate objective estimates of the probabilities of the different outcomes and those who have access to privileged information that is used to form such estimates. This distinction has rarely been made in the literature, apart from Peirson and Smith (2010) and Coleman and McGrath (2005). The two types of gamblers are called *experts* and *insiders*.

The distinction between these two types of gambling requires further investigation. An informed gambler of either type acquires and processes information to give accurate objective estimates of the probabilities of the different possible outcomes of an event. An expert's advantage over a casual or leisure bettor is in their superior ability to process information. Additionally, it is suggested that an expert has superior strategies for determining how much to gamble when potentially profitable betting opportunities are revealed by their superior processing abilities.

By contrast, an insider has access to privileged information that allows them to form more accurate probability estimates. Examples of such information are injuries or health problems of horses or members of a sports team.

The information obtained by an insider may often be easy to interpret and not require the complex information-processing skills of an expert. By comparison, the information-processing skills of an expert are not available to most other gamblers. In the later discussion of the legality of insider and expert gambling, this important distinction is investigated in the context of the property rights to the ownership of private information and information-processing skills,

3 Review of the literature on insider and expert gambling

This literature review first considers those studies that cover informed gambling on sporting events. However, nearly all such studies fail to identify important distinguishing features between the two types of gambling. Typically, most studies refer to informed gambling as being undertaken by insiders rather than recognising the possibility of gambling by experts.

Ali (1977) is perhaps the first study to have considered informed betting explicitly and suggested that most bettors are not sophisticated in the sense that the objective probabilities of outcomes are known to those betting on sports events. However, Ali suggests that bettors are rational in that 'no one prefers a bet with a smaller winning probability and the same or lower return, or with a lower return and the same or lower winning probability, to that available to him' (1977, p. 809). This assumption allows Ali to explain the data on horse-race betting at selected American racetracks. Sophisticated gamblers have objective information on the probabilities of an event occurring and are better informed than other gamblers. However, no distinction is made between the different possible sources of this better information and, thus, differences between expert and insider gambling are not considered.

Coleman (2004) suggests that there are two types of gambler: skilled risk-averse bettors who make a profit and uninformed bettors who are risk-averse and lose. This suggestion was supported by use of data from over 20 past studies for a number of different countries.

The clearest statement of the two types of informed gamblers is to be found in Colemann and McGrath (2005). The expert gambler 'uses publicly available information, but processes it with superior skill' (p. 232), and the insider 'accesses information that is not generally available to generate superior financial returns' (p. 232). They noted that 'it is not possible to separate skill from insider knowledge using market-level data' (p. 232). The analysis of Coleman and McGrath assumed that all of the bias in Australian parimutuel horse-racing odds is caused by insiders, and they assumed that the objective probability of a horse winning can be estimated from the win parimutuel market data.[1] From this evidence, they suggested that insiders make approximately 10 per cent of bets in win markets, while their share of exotic bets is 2 per cent or less. Interestingly, for all types of bets, the insider share was much higher for horses with low odds. This result conflicts with the opposite assumption of Shin (1991, 1992, 1993).

Hurley and McDonough (1985) assumed that informed bettors in a parimutuel sports betting market know the true probabilities of horses winning and have acquired this information at a cost, while uninformed bettors have no knowledge of the probabilities. Informed bettors were further assumed to respond to the actions of their uninformed counterparts and pursue a symmetric Nash game. The implication of their hypothesis is that the bias increases with the proportion of uninformed bettors in the market, as they bet disproportionately on the longshot. The favourite–longshot bias is determined by the track take (which is regarded as a transaction cost) and the cost of acquiring race-specific information to evaluate the true probabilities. Unfortunately, their experimental evidence is not completely compatible with their information and transaction cost explanation of the favourite–longshot bias.

Sobel and Raines (2003) used a model where, in the absence of any information held by the parimutuel betting public regarding race outcomes, the expected proportion of public bets made on each runner in a parimutuel market is $1/n$, where n is the number of race entrants and there is an extreme favourite–longshot bias. They assumed 'serious' bettors acquire race-specific information to inform their assessment of the true chances of individual runners. For the serious bettors, the degree of bias departs from the above extreme and the amounts bet approach the distribution of objective probabilities. The overall degree of bias is determined by the amount of information available to bettors, the proportion of serious bettors and the number of runners in the race. Using a substantial data set of American greyhound-racing parimutuel prices, Sobel and Raines found evidence of a conventional favourite–longshot bias associated with a high proportion of casual bettors, and of an opposite favourite–longshot bias in the presence of a high proportion of 'serious' bettors. This evidence is fully compatible with their elegant information model. However, their model and investigation does not consider the possibility of insider gambling, which is an equally possible explanation of their empirical results.

Terrell and Farmer (1996) considered a similar information model to that of Sobel and Raines (2003), in which informed bettors purchase the true probabilities of an event occurring against a background of uninformed bettors who are not able to purchase such information. The latter's betting can cause profitable bets to occur, which are taken up by the informed bettors who are assumed to be expected profit maximisers. The takeout of the track and information costs lead to the pattern of a favourite–longshot bias. The analysis of an American greyhound track suggested that 3.7 per cent of each dollar bet accrues to informed bettors, who are termed *professionals*. The professionals were observed to be a separate group using computers and processing information from past races. They were attracted by large bet pools, and many earned their livelihood through gambling. The operators at the track suggested that about 10 per cent of the bet pool comes from the wagers of these informed gamblers. Thus, these gamblers are clearly considered to be experts and not insiders.

Smith *et al.* (2006) further substantiated the information-based explanation of bias in a comparative study of betting exchange and bookmaker markets, in which they found bias to be positively related to transactions costs and negatively related to the amount of race-specific information available to the general public. Again, asymmetric information between the general betting public and informed gamblers leaves the question open as to whether the latter are insiders, experts or a mix of the two.

The information-based models of betting behaviour reviewed here implicitly consider skilled bettors. However, the acquisition of information at a cost can equally well refer to payment for inside information from insider sources. Alternatively, one might consider the low remuneration of staff with access to inside information as a cost of obtaining inside information. Thus, information-based models of sports betting could explain both gambling by insiders and experts. We now turn to consider those models that more explicitly refer to gambling by insiders.

Probably the most referenced and important models of the consequences of insider gambling are those developed by Shin (1991,1992, 1993) to explain bookmaker odds for UK horse races. The models explain the favourite–longshot bias observed in bookmaking markets as a result of bookmakers' reactions to asymmetric information, where insiders know the outcome of a race. The bookmaker response is modelled as an adverse selection problem, with the empirical consequence that bookmaker odds on longshots as a class are reduced below fair odds to prevent losses in the face of the possibility of insider bets. This action of bookmakers protects them against gamblers with privileged insider information. In the Shin model, this exposure to uncertainty is greater for low probability horses. Thus, this effect of potential insider information and gambling falls in magnitude as the expected probability of winning increases, and can consequently explain the favourite–longshot bias.

Schnytzer and Shilony (2003) adapted Shin's stylised model to allow for outsider bets to be biased. They showed that in the UK horse-racing gambling market this bias exists and can explain the favourite–longshot bias without the need for the existence of insiders, and empirically the two effects can be used to explain the existence of the bias.

A large number of studies have estimated the level of insider trading using Shin-type models. This estimation makes use of the relation between the sum of bookmaker prices (the overround) and the number of runners in a race depending on the proportion of insider betting, usually termed z. Using a large data set, Vaughan Williams and Paton (1997) find the degree of insider gambling to be in the region of 2 per cent, a figure very similar to that found by Cain *et al.* (1999) and Shin (1993).

In relation to horse racing, Vaughan Williams and Paton (1997) found that the favourite–longshot bias was more pronounced in low-grade races than in high-class races. The proportion z of insider gambling in such races was explained by there being less public and media scrutiny of low-grade races and runners. However, the ability of expert gamblers to detect profitable bets in such races is also likely to be greater. Consequently, a greater proportion of expert gambling on lower-grade races is an equally plausible explanation of these empirical results.

Cain *et al.* (2003) estimated the z degree of insider trading in bookmaker gambling markets for a range of sports. The estimates of the percentage of insiders in the markets are: baseball – 2 per cent; boxing – 6 per cent; cricket – 8 per cent; greyhounds – 4 per cent; horse racing – 3 per cent; soccer – 5 per cent; snooker – 7 per cent; and tennis – 8 per cent. However, the highest individual event levels of insider gambling were found for greyhound and horse racing. Cain *et al.* explained the lack of insider gambling in baseball by Woodland and Woodland's (1994) view that baseball has the most knowledgeable gamblers and that there is little or no favourite–longshot bias in baseball betting. Additionally, the Shin model attributes all favourite–longshot bias to insider activity and, thus, with no bias there can be no insider gambling. Paton *et al.* (1999) estimated z for fixed odds and spread betting in the UK's Premier League for football. The estimates were 3 per cent and 1.5 per cent, respectively, with the lower estimate for spread betting being explained by the tighter regulation of UK spread betting.

Cain *et al.* (2001b) found the z measure of insider trading to be positively linked to the degree to which parimutuel returns on UK horse races exceed those given by bookmakers. This difference can be explained by informed gamblers wishing to place bets at fixed odds with bookmakers, rather than in the parimutuel market where placing a bet reduces the winning return. Again, this is evidence of informed gambling rather than specifically the effect of insiders using private information to make winning bets.

The models of insider trading considered to this point assume that all bets are made simultaneously, with gamblers aware of the *ex post* returns to bets on different events. However, both parimutuel and bookmaking gambling markets occur across time, and informed betting may be observed when the odds on an outcome drop (plunge) dramatically. This effect is often interpreted as evidence of insider gambling and has been the subject of theoretical and empirical investigation. Most of the studies containing this effect that are reviewed below assume that any observed effect is the result of insider gambling. However, most of the models and evidence could equally well be explained by expert rather than insider gambling.

Dowie (1976), using UK horse-racing data, found that the correlation between realised probabilities and the initial odds (the opening prices) in the bookmaking markets was not significantly less than the correlation between the realised probabilities and the final odds (the starting prices). It might be expected that insider gambling would increase the correlation. However, if the insider gambling is only a small proportion of total gambling, its effect may be heavily masked by the much larger amount of uninformed betting.

Crafts (1985) argued that if the odds on a horse decline markedly in the bookmaking market, this is evidence of insider trading. Crafts found that for UK horse races, betting on such horses at the forecast price in the morning media before the bookmaking market opened would be profitable. A later paper by Crafts (1994) showed that for horses that had not run for a long time and showed a marked fall in the forecast price to the starting price, a profitable return could be made at both sets of prices. The complete lack of previous race form or other public information indicates that this is strong evidence of insider (and not expert) trading. However, over a five-year period, only 88 horses were in this category. Thus, though the rates of return were high, this is a small number of horses (in the 1973 Flat season there were 29,307 runners in 2,777 races).

By contrast, Bird and McCrae (1987) found that movements in bookmaker odds at three Australian racetracks could not be used as a basis of a profitable gambling strategy, whatever time bets were placed at.

Tuckwell (1983) argued that Australian bookmakers' margins were much less than their theoretical margin at starting prices because, on average, they accepted bets on winning horses at odds above starting prices. As bookmakers were regarded 'as keen judges of … form' (p. 112), only a small fraction of bettors are 'professionals' and most public information was available at the start of the betting market, Tuckwell argued that the cause of this lack of profitability was the betting by insiders.

The degree to which bookmakers respond to informed betting by adjusting bookmaker odds was investigated by Peirson and Blackburn (2003). They found that, for UK horse races, there was no statistically significant difference between the adjustment in bookmaker odds for winning and losing horses and the magnitude of the differences were small for all odds categories. This suggests that if bookmakers respond to insider betting, it is limited to a few horses.

Schnytzer and Shilony (1995), in a novel use of a natural experiment, compared on-course and off-course betting in Australian betting markets. Bookmakers only operate at Australian race courses; the off-course betting is catered for by a parimutuel system that is also available to those at the race course. On-course bettors observing a sudden plunge in the odds offered by bookmakers on a horse may take this as being evidence of informed betting and, although the plunge may remove the possibility of a profitable bet with a bookmaker, the opportunity would exist for the on-course gambler to bet on the parimutuel. Through comparing on-course parimutuel gamblers with off-course gamblers, who have no access to this plunge information, Schnytzer and Shilony argued that the plunges were caused by valuable inside information.

Law and Peel (2002) suggested that the cause of plunges in odds at UK horse races may be insider activity or uninformed herd-like behaviour. They found that when the Shin *z* measure of insider activity declines with the plunge, the profitability of betting at the final odds was negative, and this indicated that the cause is herd behaviour. Conversely, if the Shin measure increased, this was associated with positive returns from betting at the final odds. Schnytzer and Snir (2008) showed that there is strong evidence of herding behaviour in the UK and Australian on-course bookmaker horse-race betting markets.

Ottavani and Sorenson (2003, 2006, 2009, 2010) use a theoretical model that informed bettors have an incentive to place bets late in a parimutuel betting market as such betting gives insufficient time for other, less informed, gamblers to take advantage. It is shown that such behaviour can be used to explain the favourite–longshot bias. This tendency had been noted by Asch *et al.* (1982). However, it should be noted that in an experiment by Camerer (1998), it was found that moderately large parimutuel bets did not have a systematic effect in moving the later odds. However, Camerer noted that the market may ignore the impact of one largish bet unless it is followed by a steady flow of bets.

Shin-type models assume that an insider knows the outcome of a race and that outsiders bet on a horse in proportion to the probability of the horse winning. In both cases, these assumptions are technically useful but unrealistic (see Schnytzer *et al.*, 2008). Schnytzer and Shilony (2005) and Schnytzer *et al.* (2008) take account of real-time betting and variation in bookmaker prices in a Shin-type model of a bookmaker who faces insider gambling. In a complex call option financial model using Monte Carlo simulations, Australian bookmakers' evidence on odds plunges and variation in odds, Schnytzer *et al.* suggest that insider trading is at least 20 per cent of the Australian racetrack betting market.

Marginson (2009) argues that the Shin measure of insider trading can strictly only be applied to bookmaking markets as the measure results from bookmakers protecting themselves from insiders placing bets on outcomes that they know will occur because of their superior information. Thus, Marginson criticises Smith *et al.*'s (2006) estimation of the Shin measure for betting exchanges and, by extension, parimutuel betting markets. Marginson suggests that the exchanges give new opportunities for insiders to lay horses to lose. The extent of insider gambling on exchanges is difficult to quantify. Laying on betting exchanges may be a significant problem for sports betting involving human participants who may have some injury, problem or characteristic unknown to the general public.

Coleman (2007) estimated Shin's z measure of insider activity for the Australian horse-race bookmaker market at 2 per cent. However, the estimate for the parimutuel market is nearly identical, though Shin's model would suggest a zero value because in this model the favourite–longshot bias is purely determined by the behaviour of the bookmaker. Coleman warns against the use of the Shin measure of insider activity in parimutuel markets as an example of the 'joint hypothesis problem' in which the market model is mis-specified.

Peirson and Smith (2010) consider unraced two-year-old horses as the most likely type of racehorse to be associated with inside information. For UK bookmaker data, they found that these horses were 16 per cent less likely to win than odds-matched previously raced two-year-old horses. This effect was more pronounced for outsiders than for horses likely to win. This is evidence of insider gambling and explains some, but by no means all, of the favourite–longshot bias. These results suggest that insider information exists but is not the only type of informed gambling.

Models of expert and insider gamblers are difficult to distinguish, and the impact of the two types of gambler on sports betting markets prices are very similar. Thus, models that attribute the cause of bias in odds entirely to gambling by insiders are incorrect. The cause of bias could equally be the response of bookmakers or parimutuel markets to expert gamblers processing the publicly available information or, more likely, both types of gamblers could cause the biases. Additionally, there may well be other causes of the favourite–longshot bias. Thus, there are reasons why past studies may have overestimated the impact on odds of gambling by insiders. Policymakers need to be aware of this overestimation when framing regulations to deal with insider trading in such markets. The conclusion that there is an implicit confusion in the modelling of the characteristics and impacts of gambling by insiders and experts leads to the need for models that can distinguish between the two types of gambling. The next two sections consider these issues.

4 Characteristics of expert and insider gambling

The previous literature review suggests that there may be few, if any, differences between expert and insider gambling. Coleman and McGrath (2005) suggest that 'it is not possible to separate [gambling] skill from inside knowledge using market-level data' (p. 232). In this section, the possibility of identifying observable and distinctive characteristics of the two types of gambling is investigated.

It is suggested that expert gamblers may, on balance, tend to bet on outcomes that have a high probability of occurring. It would be expected that expert gamblers approach the task of gambling for profit in a rational manner and their behaviour should be capable of being represented by risk-averse expected utility functions. Peirson (2010) showed that risk-averse agents, who are expected utility maximisers and face different gambling opportunities with the same relative bias in their favour, will prefer to bet on the outcomes with higher probabilities. Additionally, they will choose to bet more heavily on such outcomes to the extent that their winnings on such bets would exceed winning bets on events with lower probabilities of success. The relevance of this theoretical result to the present investigation requires further assumptions. It is assumed that expert gamblers are equally skilled in finding profitable gambling opportunities across the probability range, and the distribution of the implied bias in their favour is assumed not to vary across probabilities. Thus, expert gamblers are more likely to gamble on high-probability events. This suggestion would lead to a reverse favourite–longshot bias from the gambling by experts. It could be used to explain the reverse bias found in Hong Kong horse-race betting markets, where it may be the case that gamblers behave more like experts (see Busche and Hall, 1988). Empirical evidence compatible with a reverse favourite–longshot bias for expert and informed betting away from the track and on weekdays has been found by Sobel and Raines (2003), Bruce *et al.* (2009, 2010) and Sung *et al.* (2010).

By comparison, gamblers using inside information and making no use of any further decision-making skills may be characterised as behaving in a manner similar to the general population. Thus, insiders may be expected to show a relative preference for gambling on low-probability events. The preference in the general population for low-probability events has been documented by Tversky and Khanemann (1992), Khanemann and Tversky (1979, 1982) and many others. In the gambling literature, this tendency has been discussed extensively (Vaughan Williams, 1999). That a preference for gambling on low-probability outcomes results in a standard favourite–longshot bias is demonstrated by Shin (1991, 1992, 1993) in the case of odds set by bookmakers, and, by Hurley and McDonough (1985) and Sobel and Raines (2003) in the case of parimutuel markets.

In the above analysis, we consider the consequences of the different preferences with regard to risk and probability of winning of expert and insider gamblers. However, the two types of gamblers are likely to be presented with different types of gambling opportunities. Inside information may occur for rare and highly probable events. However, in sports betting we might predict that the degree of media attention is greater for horses, teams, etc. that are more likely to win. Thus, the likelihood of inside information remaining privileged is less. Thus, as suggested by Shin, the possibility of inside information existing decreases with the probability of the event increasing. The possibility of inside information existing may increase as the degree of public interest and attention is smaller. Thus, Vaughan Williams and Paton (1997) found that, in lower-grade handicap races, the degree of insider gambling is greater, which they attributed to the lesser attention paid by the public and media to such races.

By comparison, expert gamblers use their decision-making skills to identify opportunities for profitable gambles. There are two important effects determining the opportunities found by expert gamblers. First, the degree of media attention on horses, teams, etc. that are more likely to win is greater. The greater volume of information and detailed public discussion of the likelihood of success suggest that experts will find it more difficult to detect profitable gambling opportunities on events with a high probability of occurring. Second, there is a preference of uninformed (e.g. casual or leisure) gamblers to place bets on low-probability events, and they have a general reluctance to bet on high-probability events. Such behaviour results in a favourite–longshot bias that is likely to present relatively more profitable gambling opportunities on high rather than low-probability events. The latter effect may dominate. Evidence of gambling on weekday and weekend greyhound and horse races in America and the United Kingdom suggest that there are greater opportunities for informed gambling at the weekend when there are more leisure and uninformed gamblers placing bets (Sobel and Raines, 2003; Sung *et al.*, 2010).

This limited analysis suggests tentatively that relatively experts are more likely to gamble on high probability events and insiders more likely to gamble on low probability events. This would appear to be the only possible observable distinguishing feature between insider and expert gamblers.

5 The legality of expert and insider gambling

Gambling by experts is usually considered to be a completely legal activity and is often applauded with an element of envy and notoriety (Veitch, 2009). Profitable betting by experts is usually regarded as an appropriate reward for the development and ownership of superior information-processing skills.

By contrast, gambling on sporting events by those with inside information is often considered to be highly undesirable and widely condemned (Forrest and Simmons, 2003). However, a clear economic analysis of the costs and benefits of insider betting is not available. To partly fill this omission, the literature on insider trading in financial securities is reviewed and applied to the case of gambling on sports events.

The standard ethical arguments against insider trading are that it is unfair, involves the misappropriation of information, harms ordinary investors and distorts the market (Moore, 1990). The unfairness argument follows from the parties to the transaction having unequal information or unequal access to information. It is suggested that the unfairness follows from a potential fiduciary duty to reveal information. Fairness of equal information for the two parties involved in a transaction depends on how information advantages are acquired and whether the party with superior information is entitled to profit from the advantage. Crafts (1985) suggested that insider gambling on horse races is an acceptable reward for those involved in owning and training horses and can be regarded as compensation for the high costs for owners and trainers, and low wages of their staff.

Information can be regarded as a property, and the use of inside information is potentially theft from the owner. In the case of sports betting, ownership must surely lie with the owners of the team, horse, etc., or the organisation running the sporting competition. Thus, if insider trading is mainly a problem of misappropriation of the owner's information, then it should only be illegal when the owner prohibits the use of this information by others – for example, her staff. In particular, in these circumstances, there cannot be a reason to restrict the owner's use of inside information. However, if the ownership of information is considered to lie with the organisation running the sporting event, insider gambling may present more ethical problems. The organisation will be concerned that competitors are seen to try their utmost. This endeavour hopefully ensures a greater sporting spectacle and confidence in betting on the event. The possibility that competitors may not try or may cheat is considered next.

It is suggested that insider trading will lead to the pursuit of activities that lead to inside information, and these activities may be socially wasteful or harmful. Thus, a horse-race trainer may dope a horse or deliberately instruct jockeys to not try to win races in order to improve the odds available in future races when the horse is raced to win. Alternatively, bookmakers may seek to obtain information from the participants in a sporting event to help them set odds, or they may offer bribes or make threats to obtain information. Thus, insider gambling may lead to more cheating or failing to try in sports events which attract large gambling markets. For the organisers of sporting events, the integrity of the sport may be jeopardised by insider gambling.

The emergence of internet betting has vastly expanded the size of gambling markets on many sporting events. The internet allows the placing of bets on sporting events thousands of miles away. However, in terms of regulation, the internet potentially allows tracking of gamblers and the volume of betting support for the occurrence of unlikely sporting outcomes. So, as strongly argued by the owners of betting exchanges,[2] tracking of all bets on sporting events potentially allows for the early recognition of possible corrupt outcomes. Perhaps the most important aspect of internet gambling is the new possibility of betting on events not to occur – this is called *laying*. For example, one can bet on a horse not to win a race. The possibility of laying makes the use of inside information potentially more profitable and gives greater incentives for outright cheating.

Insider trading may harm the other participants in a market. In the case of parimutuel sports betting, outsiders may bet on an outcome whose probability of occurring is much less than that implied by the betting market prices. Thus, the outsiders are victims and most likely are unaware of this crime. In the case of sports betting with bookmakers who set prices of bets, these prices will, in general, reflect the possibility of insider information and make betting less attractive for the outsider.

It is relevant to consider the arguments that have been made in support of insider trading and how they may apply to insider gambling. The major argument in favour of insider trading is that it results in greater market price efficiency (Manne, 1966; Bainbridge, 2001). In the case of parimutuel and bookmaking gambling, the previous review shows that the presence of informed betting is likely to lead to more efficient prices as parimutuel markets adjust to the heavy betting by insiders, or bookmakers adjust their odds when receiving bets from traditionally successful bettors. Presumably, efficient prices are important and desirable in betting markets. However, as compared to financial markets, the desirability of efficient prices to the leisure gambler is not perhaps as great. In addition, and as noted by Kay (1988), regulation against insider activity favours market professionals as it opens more opportunities for making profitable bets.

An additional argument in favour of insider trading is that it provides 'the most appropriate device for compensating entrepreneurs in large corporations' (Manne, 1966, p. 182). As mentioned previously, this argument is used by Crafts (1985) to justify insider gambling by the horse-racing industry to cover the large costs of training and the low staff wages. As such, this constitutes a second-best argument rather than placing a (increased) levy on the gambling industry to subsidise the horse-racing industry that operates at a net aggregate loss.

No attempt is made to come to a conclusion about the ethical status of insider gambling, but, as noted by Moore (1990) in the context of insider trading, 'this practice has received surprisingly little ethical analysis' (p. 171).

6 Conclusion

This survey shows that the literature on sports betting by insiders and experts fails to make the important distinction between the two types of betting. The literature tends to focus on the effects of insider gambling, though the survey shows that the theoretical and empirical models could be used equally well to explain gambling by experts. Empirical estimates of the insider proportion of total betting vary between 0 and 20 per cent. These estimates, though widely quoted, are in doubt as the underlying models do not distinguish between expert and insider gambling. It is suggested that future empirical models should examine sports betting where there is more or less likely to be insider or expert gambling. An example of this would be where a horse has not previously raced, either at all or in this season, as suggested by Peirson and Smith (2009) and Crafts (1994). Alternatively, one can consider sports events where there is likely to be a greater concentration of experts, such as the weekday greyhound races considered by Sobel and Raines (2003). Use may be made of the suggestion that expert gamblers may bet relatively more frequently on high-probability outcomes than insiders.

The distinction between insider and expert gambling is important as the former is usually regarded as wrong and often is illegal. However, in spite of the current concern about the impact of gambling on the integrity of sporting events, there is no clear theoretical analysis of the ethics and fairness of insider gambling, and this omission is important. This survey uses the literature on insider trading to develop such an analysis. However, the applied ethics of insider gambling requires value judgement assumptions to be made on the ownership of the insider information; the fairness of gambling with asymmetric information; the effects of insider gambling on the efficiency of the prices offered by bookmaker and parimutuel markets; and the impact of insider gambling on sports cheating.

Notes

1 The statistical analysis does not test for serial correlation or specification error, which is important in the modelling of the non-linear relation between objective and subjective probabilities in a betting pool market.
2 For example, see the views of Betfair at http://corporate.betfair.com/media/integrity-in-sports-betting.html.

References

Ali, M.M. (1977), 'Probability and Utility Estimates for Racetrack Bettors', *Journal of Political Economy*, 85, 803–815.

Asch, P., Malkeil, B.G. and Quandt, R. (1982), 'Racetrack Betting and Informed Behavior', *Journal of Financial Economics*, 10, 187–194.

Bainbridge, S.M. (2001), *The Law and Economics of Insider Trading: A Comprehensive Primer*, paper, School of Law, University of California, UCLA.

Beny, L.N. (2002), *The Political Economy of Insider Trading Legislation and Enforcement: International Evidence*, Discussion Paper 348, Harvard Law School, Cambridge, MA.

Bird, R. and McCrae, M. (1987) 'Tests of the Efficiency of Racetrack Betting Using Bookmaker Odds', *Management Science*, 33 (12), 1552–1562.

Bruce, A., Johnson, J., Peirson, J. and Yu, J. (2009), 'An Examination of the Determinants of Biased Behaviour in a Market for State Contingent Claims', *Economica*, 4 (1), 282–303.

Bruce, A., Johnson, J., Peirson, J. and Yu, J. (2010), 'A Decision Making Perspective on Home Asset Bias and Investment Performance in a Financial Market', paper, Centre for Risk Research, University of Southampton.

Busche, K. and Hall, C.D. (1988), 'An Exception to the Risk Preference Anomaly', *Journal of Business*, 61, 337–346.

Cain, M., Law, D. and Peel, D.A. (1999), 'Estimates of the Degree of Insider Trading in two Disparate Betting Markets', *Applied Economics Letters*, 6 (3), 191–193.

Cain, M., Law, D. and Peel, D. (2001a), 'The Relationship between Two Indicators of Insider Trading in British Racetrack Betting', *Economica*, 68, 97–104.

Cain, M., Law, D. and Peel, D. (2001b), 'The Incidence of Insider Trading in Betting Markets and the Gabriel and Marsden Anomaly', *The Manchester School*, 69, 197–207.

Cain, M., Law, D. and Peel, D. (2003), 'The Favourite-Longshot Bias, Bookmaker Margins and Insider Trading in a Variety of Betting Markets', *Bulletin of Economic Research*, 55, 263–273.

Camerer, C. (1998), 'Can Asset Markets Be Manipulated? A Field Experiment with Racetrack Betting', *Journal of Political Economy*, 106, 457–482.

Clotfleter, C.T. (2005), 'Gambling Taxes', in S. Cnossen (ed.), *Theory and Practice of Excise Taxation: Smoking, Drinking, Gambling, Polluting and Driving*, Oxford University Press: Oxford.

Coleman, L. (2004), 'New Light on the Longshot Bias', *Applied Economics*, 36, 315–326.

Coleman, L. (2007), 'Just How Serious is Insider Trading? An Evaluation using Thoroughbred Wagering Markets', *Journal of Gambling Business and Economics*, 1, 31–55.

Coleman, L. and McGrath, M. (2005), 'Biases and Insider Trading in Exotic Bets on Thoroughbreds', in L. Vaughan Williams (ed.), *Information Efficiency in Financial and Betting Markets*, Cambridge University Press: Cambridge.

Crafts, N.F.R. (1985), 'Some Evidence of Insider Knowledge in Horse Race Betting in Britain', *Economica*, 52, 295–304.

Crafts, N.F.R. (1994), 'Winning Systems? Some Further Evidence on Insiders and Outsiders in British Horse Race Betting', in D.B. Hausch, V. Lo and W.T. Ziemba (eds), *Efficiency of the Racetrack Betting Market*, Academic Press: New York.

Dowie, J. (1976) 'On the Efficiency and Equity of Betting Markets', *Economica*, 43, 139–150.

Forrest, D. and Simmons, R. (2003), 'Sport and Gambling', *Oxford Review of Economic Policy*, 19, 598–611.

Hurley, W. and McDonough, L. (1995) 'A Note on the Hayek Hypothesis and the Favourite-Longshot Bias in Parimutuel Betting', *American Economic Review*, 85 (4), 949–955.

Kahneman, D. and Tversky, A. (1979), 'Prospect Theory: An Analysis of Decision Under Risk', *Econometrica*, 47, 263–291.

Kahneman, D. and Tversky, A. (1982) 'Intuitive Prediction: Biases and Corrective Procedures', in A. Tversky, D. Kahneman and P. Slovic (eds), *Judgment Under Uncertainty: Heuristics and Biases*, Cambridge University Press: Cambridge.

Kay, J. (1988), 'Discussion', *Economic Policy*, 3, 163–193.

Law, D. and Peel, D.A. (2002) 'Insider Trading, Herding Behaviour, and Market Plungers in the British Horse Race Betting Market', *Economica*, 69, 327–338.

Manne, H.G. (1966), *Insider Trading and the Stock Market*, The Free Press: New York.

Marginson, D. (2009), 'Market *In*efficiency in Person-to-Person Betting: Examining "Known Loser" Insider Trading on the Exchanges', *Southern Economic Journal*, 76, 932–952.

Moore, J. (1990), 'What is Really Unethical about Insider Trading?', *Journal of Business Ethics*, 9, 171–182.

Ottaviani, M. and Sorenson, P.N. (2003), 'Late Informed Betting and the Favourite-Longshot Bias', CEPR Discussion Paper 4092, London.

Ottaviani, M. and Sorenson, P.N. (2006), 'The Timing of Parimutuel Bets', paper, London Business School. www.econ.ku.dk/sorensen/Papers/pvfom.pdf.

Ottaviani, M. and Sorenson, P.N. (2009), 'Aggregation of Information and Beliefs: Asset Pricing Lessons from Prediction Markets', University of Copenhagen Department of Economics Discussion Paper 09–14.

Ottaviani, M. and Sorenson, P.N. (2010), 'Noise, Information and the Favorite-Longshot Bias in Parimutuel Predictions', *The American Economic Journal: Microeconomic*, 2, 58–85.

Paton, D., Vaughan Williams, L. and Fraser, S. (1999), 'Regulating Insider Trading in Betting Markets', *Bulletin of Economic Research*, 51 (3), 237–241.

Peirson, J. (2010), 'Why does Expected Utility Favour High probability Events?', paper, School of Economics, University of Kent.

Peirson, J. and Blackburn, P. (2003), 'Betting at British Racecourses: A Comparison of the Efficiency of Betting with Bookmakers and at the Tote', in L. Vaughan Williams (ed.), *Economics of Gambling*, Routledge: London.

Peirson, J. and Smith, M. (2010), 'Expert Analysis and Insider Information in Horse Race Betting: The Danger of Regulating Virtuous Market Behavior', *Southern Economic Journal*, 76 (4), 976–992.

Sauer, R.D. (1998), 'The Economics of Wagering Markets', *Journal of Economic Literature*, 36, 2021–2064.

Schnytzer, A. and Shilony, Y. (1995), 'Inside-Information in a Betting Market', *Economic Journal*, 105, 963–971.

Schnytzer, A. and Shilony, Y. (2003), 'Is the Presence of Insider Trading Necessary to Give Rise to a Favourite-Longshot Bias?', in L. Vaughan Williams (ed.), *The Economics of Gambling*, Routledge: London.

Schnytzer, A. and Shilony, Y. (2005), 'Insider Trading and Bias in a Market for State-Contingent Claims', in L. Vaughan Williams (ed.), *Information Efficiency in Financial and Betting Markets*, Cambridge University Press: Cambridge.

Schnytzer, A. and Snir, A. (2008), 'Herding in Imperfect Betting Markets with Inside Traders', *Journal of Gambling Business and Economics*, 2, 1–15.

Schnytzer, A., Lamers, M. and Makropoulou, V. (2008), 'The Impact of Insider Trading on Forecasting in a Bookmakers' Horse Betting Market', paper, Department of Economics, Bar-Ilan University, Israel.

Shin, H.S. (1991), 'Optimal Betting Odds Against Insider Traders', *Economic Journal*, 101, 1179–1185.

Shin, H.S. (1992), 'Prices of State Contingent Claims with Insider Traders, and the Favourite-Longshot Bias', *Economic Journal*, 102, 426–435.

Shin, H.S. (1993), 'Measuring the Incidence of Insider Trading in a Market for State-Contingent Claims', *Economic Journal*, 103, 1141–1153.

Smith, M.A., Paton, D. and Vaughan Williams, L. (2006), 'Market Efficiency in Person-to-Person Betting', *Economica*, 73, 673–689.

Sobel, R.S. and Raines, S.T. (2003), 'An Examination of the Empirical Derivatives of the Favourite-Longshot Bias in Racetrack Betting', *Applied Economics*, 35, 371–385.

Sung, M.-C., Johnson, J. and Peirson, J. (2010) 'Understanding the Weekend Effect: Insights from a Market for State Contingent Claims', paper, Centre for Risk Research, University of Southampton.

Terrell, D. and Farmer, A. (1996), 'Optimal Betting and Efficiency in Parimutuel Betting Markets with Information Costs', *Economic Journal*, 106, 846–868.

Thaler, R. and Ziemba, W. (1988), 'Parimutuel Betting Markets: Racetracks and Lotteries', *Journal of Economic Perspectives*, 2, 161–174.

Tuckwell, R. (1983), 'The Thoroughbred Gambling Market: Efficiency, Equity and Related Issues', *Australian Economic Papers*, 22, 106–108.

Tversky, A. and Kahneman, D. (1992), 'Advances in Prospect Theory: Cumulative Representation of Uncertainty', *Journal of Risk and Uncertainty*, 5, 297–323.

Vaughan Williams, L. and Paton, D. (1997), 'Why is there a Favourite-Longshot Bias in British Racetrack Betting Markets?', *Economic Journal*, 107, 150–158.

Vaughan Williams, L. (1999), 'Information Efficiency in Betting Markets', *Bulletin of Economic Research*, 51, 1–30.

Veitch, P. (2009), *Enemy Number One: The Secrets of the UK's Most Feared Professional Punter*, Highdown: London.

Woodland, L.M. and Woodland, B.M. (1994), 'Market Efficiency and the Favourite-Longshot Bias: The Baseball Betting Market', *Journal of Finance*, 49, 269–279.

15 Who can beat the odds?

The case of football betting reviewed

Anastasios Oikonomidis and Johnnie Johnson

1 Historical development of football betting markets

Sports betting has been an 'integral part of working class structure' in the United Kingdom since the beginning of the twentieth century (Jones *et al.*, 2000). At the end of the 1960s, nearly a decade after the formalization of sports betting markets (betting shops were legalized in the United Kingdom in the early 1960s), nearly 16,000 betting shops operated in the United Kingdom, though concentration resulted in this declining to about 8,800 in 1998 (Jones *et al.*, 2000). However, betting turnover has been increasing steadily, leading the four major British bookmaking firms to report turnover of £10 billion in aggregate, in 2002 (Levitt, 2004). Global Betting and Gaming Consultants (2001) indicated that in 1998 about four million adults were betting weekly on sports in the United Kingdom (Forrest and Simmons, 2003). More specifically, football is, according to Mintel (2001), the fastest-growing form of gambling in the United Kingdom, and the Gambling Review Report (DCMS, 2001, paragraph. 9.23) indicates that most sports betting activity concerns football (Forrest and Simmons, 2003).

Obviously, the popularity of sports betting is not constrained to the United Kingdom. Worldwide, betting on sports is extremely popular and football betting has the lion's share in most countries. The National Gambling Impact Study Commission (1999) estimates that wagering on sporting events in the United States approaches US$380 billion annually (Levitt, 2004). In 2007/2008, The Jockey Club was the greatest Hong Kong tax payer, contributing about US$1.690 billion, approximately 6.5 per cent of all taxes collected by the Hong Kong government (Wong *et al.*, 2009). Similar data are reported by So and Kwok (2007) for the 2005/2006 season. They also show that football betting turnover in the 2009/2010 season was HK$31.27 billion (about £2.67 billion). Moreover, according to Forrest and Simmons (2003), football industry sources suggest that Far Eastern football betting turnover is about US$1 billion per weekend during the course of the football season, and that about half of this is bet on English Premier League matches.

During the last decade, due to the spread of the internet, online betting has developed rapidly. According to Jones *et al.* (2000), Sportingbet, the then small company established by an independent bookmaker, was the first company to enter the internet sports betting market. Many established companies followed and new ones were formed, creating a very competitive and dynamic market setting. Malaric *et al.* (2008) identify 600 different sports betting web sites operating worldwide, representing a US$16.6 billion market, which is predicted to rise. The number of internet gambling companies (including casino operations), has been reported to exceed 1,800 (Forrest and Simmons, 2003), and the Sport-Business Group (2001) predict that the e-gaming industry will exceed US$100 billion by 2015. China, Hong Kong, Singapore and Sweden are named as countries which offer high prospects for growth (Forrest and Simmons, 2003).

Another key feature associated with the evolution of sports betting markets is the establishment of betting exchanges. The function of betting exchanges resembles that of the honest brokers in the eighteenth century. Hence, unlike the bookmakers, the betting exchange company acts as an intermediary that matches opposing bets between punters, holds the funds until the outcome is decided and pays the winner, after deducting a small commission. This commission, being risk-free, allows the company to set it at a lower level compared to the bookmakers' usual margin. Sporting Exchange Limited is the major betting exchange company, with the trading name Betfair. It was founded in 1999 and launched its website in June 2000 (Jones *et al.*, 2004). On a daily basis, Betfair matches about 500,000 bets and had reported turnover exceeding £50 million per week in 2003 (Jones *et al.*, 2004). The company has over two million registered users (Croxson and Reade, 2008). Even though several other betting exchange companies, such as Sporting Options, Betdaq and GGBet have entered the market (Jones *et al.*, 2004), Betfair accounts for 90 per cent of all exchange-based betting activity worldwide (Croxson and Reade, 2008).

Football betting also takes place in spread-betting, prediction and pari-mutuel markets. The development of betting exchanges and the spread of internet betting has intensified competition, forcing bookmakers to decrease their margins (known as the over-round) significantly on football bets in recent years (Forrest and Simmons, 2003; Oikonomidis and Johnson, 2008). In addition, recent developments in legislation have been beneficial to punters (e.g. regarding UK betting tax – see Paton *et al.*, 2002). Thus, overall, modern football betting markets are associated with friendly legislation in many places in the world, high volume, intense competition and very low transaction costs. Consequently, it could be claimed that these markets constitute the ideal setting for the exploitation of profitable opportunities. Testing the efficiency of the football betting market with respect to different sources of information therefore presents an interesting opportunity for researchers. In the following sections, the extent to which this has been achieved will be reviewed. This will be structured around Fama's (1970) categorization of market efficiency; consequently, literature regarding 'weak-form', 'semi-strong-form' and 'strong-form' efficiency of football betting markets will be examined in turn.

2 Weak-form efficiency

2.1 Introduction

According to Fama (1970), a market is weak-form efficient if current prices reflect all information arising from past prices. Consequently, a betting market is weak-form efficient if abnormal returns cannot be made using any kind of information related to market odds. A significant number of papers have analysed football betting markets with respect to odds information. Thus researchers have explored the value of odds in predicting football events, the existence of systematic odds-related biases, as well as the degree of variation of odds between different market operators. In the following section, these studies are reviewed.

2.2 Odds as predictors

In a football betting market, odds reflect the estimations of market makers regarding the probability of competing outcomes; some papers have explored the accuracy of such odds. For example, Leitner *et al.* (2008) analyse odds quoted by 45 bookmakers (concerning the European Championship 2008 competition) to explore how successfully these predict match outcomes. They employ mixed effects regression (group- and bookmaker-specific fixed effects and team-specific random effects) to model the true odds of each team winning the competition, based on the market odds. They compute pairwise winning probabilities and simulate the tournament, concluding that the estimated odds-based probabilities are highly correlated with the actual outcomes. Leitner *et al.* (2009) apply a similar methodology to obtain winning probabilities for teams in the Champions' League (season 2008/2009), assigning the highest probabilities to Chelsea, Manchester United, Inter-Milan and Barcelona. Three of these teams reached the semi-finals, indicating that predictions arising from the odds-based model were close to the realized outcome. Finally, data from the prediction market (STOCCER championship market) has also been analysed (Luckner and Weinhardt, 2008), and it was found that estimations arising from 'play money' were no less accurate than those arising from betting odds.

Some researchers have explored whether variation exists in the accuracy of odds-based predictions of football events. For example, Strumbelj and Sikonja (2010) examined odds of ten bookmakers related to 10,699 matches from six major European football leagues. They found that the accuracy of odds in predicting outcomes increased through time, but that variations existed in the forecasting ability of different bookmakers and in the cross-league accuracy of the odds in predicting results. However, this later finding could simply be due to cross-league differences in competitiveness. Thus, in more competitive leagues, the outcome of a football event is likely, on average, to be more random (c.f. a less competitive league).

The results of the papers discussed above suggest that football odds exhibit forecasting power; this is not unexpected given the availability of public information regarding the sport and the size of its betting market. However, in order to understand whether odds efficiently incorporate publicly available information and are therefore set at a level to prevent abnormal returns being made, an 'accuracy benchmark' is required. This benchmark can be provided by quantitative forecasting models that utilize publicly available, including fundamental information regarding football events. However, such tests move the investigation to a 'semi-strong' level and, therefore, this subject is reviewed in the corresponding section, later in this chapter.

2.3 Odds biases

Undoubtedly, the most popular object of research is the famous favourite–longshot bias (FLB). Initially, the bias was observed in horseracing, but subsequent research has documented its existence in a variety of sports betting markets (for a survey of studies see Sauer (1998) and Vaughan Williams (2005)). A FLB exists when the favourites' (longshots') winning probability, as implied by the odds, is underestimated (overestimated). However, the reverse phenomenon has also been documented (e.g. Woodland and Woodland, 1994) and is usually referred to as 'reverse' or 'negative' FLB. The literature concerning the FLB in the football betting market is now reviewed.

Pope and Peel (1988) investigated the UK market (data derived from the 1981/1982 season) and concluded that favourites seem to be more profitable compared to longshots. Cain *et al.* (2000) analysed a data set of 2,855 games played in the United Kingdom during the 1991/1992 season and found some evidence that market odds underestimate the winning probability of heavy favourites (including the probability of frequent (low) exact scores). Similar evidence was presented by Malaric *et al.* (2008), who explored a data set of 12,218 games played in ten European leagues in the period 1999–2002. Deschamps (2008) also documented FLB associated with several European leagues in the 2005/2006 season, with more pronounced effects in second-division leagues. Finally, Vlastakis *et al.* (2009) observed that the market underestimates merely the winning probabilities of favourites playing away from home.

The 'draw' outcome is more frequent in football compared to other sports and therefore it is interesting to investigate whether its existence influences the FLB in any way. Deschamps and Gergaud (2007) explore a data set of 8,377 football matches played between the 2002/2003 and 2005/2006 seasons in English leagues and observe 'positive' FLB concerning the odds of the home and away teams. However, for the odds on the draw outcome a reverse FLB is identified. Additionally, it is found that the probability for a draw is, in general, underestimated by market odds. Consequently, betting on 'draw' yields superior returns than betting on the home or the away team winning.

Variation in the magnitude of FLB has also been associated with the level of transaction costs and with league-specific characteristics, including competitiveness. Paton and Vaughan Williams (1998) found that the fixed-odds football betting market, where transaction costs are higher, exhibits higher FLB compared to the spread-betting market, where transaction costs are lower. Oikonomidis and Johnson (2008) suggested that if bookmakers fail to fully account for cross-league differences in competitiveness, heterogeneity in the magnitude of the FLB would be expected. This hypothesis was confirmed by their analysis of a sample of over 56,000 football matches played in 22 European leagues over the last decade; the level of league competitiveness almost completely determined the degree of FLB in each league, with relatively competitive leagues exhibiting significantly higher bias.

Conclusively, it can be stated that significant 'positive' FLB exists in the football betting market. The fact that it has been documented across different samples and is shown to be persistent across years points to the fact that it is a structural idiosyncrasy of the market.

2.4 Market variation

As indicated above, the football betting market is currently very large and competitive, with many companies, including bookmakers, betting exchanges and spread-betting firms, operating with low margins. Hence, it is interesting to explore to what extent odds-based information, arising from these different sources, can be used by punters to increase their returns. Relevant questions to address are: (1) whether there is sufficient variation across market prices to provide punters zero-risk opportunities to earn profit (i.e. 'arbitrage'); and (2) whether variation of market prices signals the arrival of information concerning the probabilities of particular football events, which can increase the accuracy of punters' forecasts and, as a consequence, enable them to earn abnormal returns.

Arbitrage

Pope and Peel (1988) identified arbitrage opportunities in the football betting market. However, later research (Dixon and Pope, 2004) analysing odds from three different bookmakers found no opportunities for arbitrage. Dixon and Pope suggest that this may be due to a decrease in the variation of odds between bookmakers, resulting from their prices having become more coordinated under the influence of professional arbitrageurs. Similarly, Vlastakis *et al.* (2008) explore a sample of 12,420 football matches, including odds from five different bookmakers and identify only a small number (63) of arbitrage opportunities. However, for the purpose of such analysis, it would be desirable to analyse odds across a much larger sample of bookmakers, as Oikonomidis and Johnson (2008) estimate that 'shopping' for best odds across 45 bookmakers should bring the overall over-round close to 0. In a study of odds quoted by 79 different bookmakers, Deschamps (2008) identified a relatively greater number of arbitrage opportunities (293) across the sample of 6,315 games. Similarly, Deschamps and Gergaud (2007) explored odds from several different bookmakers and found that significant price variation existed; indicating that, 'shopping' for best odds can significantly increase the punter's return.

The studies discussed above used odds from the bookmaker market only. However, Franck *et al.* (2009) analysed the possibility of arbitrage opportunities arising from simultaneously betting on outcomes of the same event in the bookmaker and the betting exchange markets. They found that the development of betting exchanges has significantly increased the frequency of arbitrage opportunities, since they have increased the variation of prices in the market. In a sample of 5,478 games, they found only ten arbitrage opportunities when considering bookmaker prices alone, but 1,450 when the analysis was extended to betting exchanges. The existence of arbitrage opportunities has also been examined in prediction markets. Luckner and Weinhardt (2008) use data from the STOCCER championship market (concerning the FIFA World Cup 2006), and found no significant evidence of arbitrage opportunities.

Signals from variation of odds

Some papers have examined to what extent cross-market variation of prices is random or whether it signals information regarding the probabilities of events. For example, Deschamps (2008) analysed data from diverse bookmakers and found that outlying odds are informative, even after considering average odds. They provided empirical evidence, which suggests that if a bookmaker is willing to offer very high odds relative to the market, this indicates that market odds are lower than they should be. However, Paton and Vaughan Williams (2005), using data from the spread-betting market, found evidence to suggest that bookmakers that offer outlying odds do not possess superior information. More specifically, they found that the average mid-point of the quoted spreads from different bookmakers is a more accurate estimation of the real outcome compared to the outlying spread. This market variation was found to be sufficient to enable profitable trading.

Previous research, discussed above, suggests that the simultaneous operation of several betting companies is likely to provide 'odds shoppers' with the opportunity to drastically decrease or even nullify transaction costs and place nearly fair (or even favourable) bets. However, it should be noted that implementing successful 'arbitraging' may involve several difficulties, which are not so obvious when theoretically examining this possibility. For example, bookmakers may change their odds or refuse to accept bets at a high level, or liquidity on the desired odds may quickly disappear from the betting exchanges. In all cases, this is a business for the fastest and most computationally efficient players (see Marshall, 2009).

2.5 Summary

Overall, the literature suggests that several types of weak-form inefficiencies exist, and it appears possible for punters to take advantage of them and to at least decrease their losses. However, in order to assert that these inefficiencies are significant and persistent enough to enable punters to achieve positive returns, more consistent evidence is required. Additionally, even if the theoretical inefficiencies were shown to exist, it remains debatable whether successful exploitation is possible.

3 Semi-strong-form efficiency

3.1 Introduction

A market is semi-strong-form efficient if market prices incorporate all relevant, publicly available information (Fama, 1970). Consequently, it should not be possible for punters to use any kind of publicly available information to estimate football event probabilities more accurately than those derived from odds; forecasting models based on fundamental information should, therefore, not lead to profitable betting strategies. A range of studies have tested the semi-strong-form efficiency of football betting markets; the methods employed in these studies and their results are reviewed below.

3.2 Forecasting methods

Several papers have estimated the winning probabilities of competing outcomes of football games, through modelling (1) the expected goals scored; (2) the goal difference; or (3) the winning outcome directly. In the first instance, count outcome regression models have been used, such as Poisson or modified Poisson (e.g. Maher, 1982; Dixon and Coles, 1997; Karlis and Ntzoufras, 2003; Dixon and Pope, 2004), negative binomial (Reep *et al.*, 1971; Baxter and Stevenson, 1988) and extreme value distributions (Greenhough *et al.*, 2002). In order to model the expected goal *difference* between two opponents, Karlis and Ntzoufras (2009) applied the Skellam's distribution. A number of researchers have employed discrete choice models (mainly ordered probit regression) in order to directly estimate the probability of competing events (Kuypers, 2000; Goddard and Asimakopoulos, 2004). Goddard (2005) performed a statistical comparison of forecasting models and found no significant difference in accuracy between models that forecast goals and those that model results directly. More recently, machine learning techniques have been applied to predict game outcomes (e.g. Vlastakis *et al.*, 2008; Strumbelj *et al.*, 2009). Finally, combinations of different types of estimation have also been considered. For example, Vlastakis *et al.* (2008) suggested encompassing techniques in order to combine forecasts from Poisson and multinomial regression models, weighted according to the accuracy of predictions.

3.3 Home advantage

Home-ground advantage plays a major role in deciding football game outcomes; home win frequency is about twice that of away wins. Crowd support, stadium familiarity and travelling are factors that have been shown to contribute to the creation of the home advantage (Courneya and Carron, 1992), as has referee bias in favour of the home side (Garicano *et al.*, 2005). However, the existence of this effect does not appear to bias market odds. Pope and Peel (1988) examined data from the 1981/1982 season and found no evidence of inefficiency regarding home advantage. Furthermore, Graham and Stott (2008), using data from the top four English leagues for 2001–2006, concluded that the home advantage is relatively constant across teams (in contrast to an earlier study by Clarke and Norman, 1995), and that bookmaker prices reflect this lack of between-teams variation in home advantage. Goddard and Thomas (2006) found that home-team advantage was underestimated by market odds in the European Championship 2004. However, the small sample size and the dependency of observations do not enable wider conclusions regarding the bias to be drawn.

3.4 Performance-measuring models

Several models incorporating a wide range of publicly available information have been employed in order to test the semi-strong-form efficiency of football betting markets. An overview of these is given below.

Dixon and Coles (1997) employed a bivariate Poisson model, whose parameters relate to home advantage and past performance (in terms of goals scored and conceded). They suggested several refinements for low scoring probabilities to fit real data more accurately, and they adjusted the likelihood function to incorporate a proximity parameter, to give more weight to recent observations. The model was fitted using English league and cup data between 1992 and 1995 and was found to yield positive returns in an out-of-sample period (1995/1996 season). Similarly, Dixon and Pope (2004) developed a Poisson model that estimated team-specific parameters concerning the ability to attack and defend (based on observed outcomes). They tested the model against bookmaker odds for the correct score and the match outright market and found evidence of market inefficiency.

Employing ordered probit regression and using data derived from matches played in England after 1987, Goddard and Asimakopoulos (2004) built a forecasting model for football results, based on a series of fundamental factors. Recent results, particularly those at home for the home team and away for the away team were identified as key forecasting factors. In addition, Goddard and Asimakopoulos (2004) found that the effect of motivation was significant, while geographical distance of travel for the away team increased home-ground advantage. Elimination from the cup competition appeared to have a negative effect on a team's subsequent league results, and teams that attracted higher attendances in their previous games were more likely to be successful in future games (controlling for other performance factors, confirming that this is not an omitted variable bias). A model combining these factors was tested against market odds and found to be profitable for high expected profit bets. Likewise, Kuypers (2000) built a model utilizing similar information, using data from the 1993/1994 and 1994/1995 seasons from the top four divisions of English football, and used the model to demonstrate some degree of inefficiency. Forrest *et al.* (2005) employed a sample of nearly 10,000 English football games over the period 1998–2003 to test a similar model; they found that their model only produced superior results to market probabilities in the early years. Thus, they suggest that the football betting market has moved towards efficiency as a result of competition between different bookmaking companies, which has forced them to improve their estimations.

3.5 Behavioural biases

Decisions in markets are made by humans and, therefore, it might be expected that biases that characterize human judgement will influence the setting of prices, and may lead to inefficiencies. A review of the studies that examine the efficiency of the football betting market in relation to behavioural biases is now presented.

It is commonly believed that casual punters behave sentimentally and place bets on the team they support; a number of researchers have examined whether such behaviour biases the odds of popular teams. Forrest and Simmons (2002) found that the winning probabilities of popular teams (i.e. defined as teams that achieve high attendance at their home games) are underestimated by market odds. Similar evidence was provided by Goddard and Asimakopoulos (2004) and Forrest and Simmons (2008), analysing data from the top Spanish (2001–2008) and Scottish leagues (2001–2005), respectively. They suggested that odds are biased in favour of popular teams, because the bookmakers try to attract sentimental bettors. Franck *et al.* (2010) documented the same effect when exploring a sample of 16,000 English football games between the 2000/2001 and 2007/2008 seasons. This effect was not apparent on weekday games. The authors suggested that this result was expected assuming that more casual, 'sentimental' gamblers bet at weekends (c.f. weekdays), thus increasing the demand for popular teams. Bookmakers, as a consequence, increase the odds for such teams in order to sustain the competition in a 'price-sensitive' market.

In order to investigate whether optimistic bias exists in the betting market, Page (2009) set the opposite hypothesis compared to the studies reviewed in the previous paragraph. The author suggested that the existence of an optimistic bias among punters would lead UK betting companies (which are more likely to have a majority of British bettors) to lower the odds for UK teams in international matches (due to the likely high demand). However, it should be noted that this would result from the optimistic bias if bookmakers balance their books, but not if they try to attract sentimental punters, as suggested by the studies reviewed above. Page (2009) analysed odds derived from 161 different betting companies for 3,585 international football matches and 5,301 European cup matches between 1998 and 2007 and found no evidence of optimistic bias. On the other hand, Bernile and Lyandres (2008) investigated returns of European football clubs traded in the stock market and found that investors overestimated their teams' expected performance, leading to abnormally negative returns. Bookmakers' odds reflect to some extent their desire to attract bets on the popular teams and prediction market prices are completely demand led. Consequently, this suggests that structural differences between prediction and bookmaker markets may lead to diverse biases in market odds.

In the total goals market, a utility bias has been observed. For example, Rodney and Weinbach (2009) analysed over 15,000 football games played in 22 European leagues. They examined the most common form of betting – i.e. to bet on whether more ('over') or fewer ('under') than 2.5 goals will be scored in the game. They found odds in this market to be significantly biased, as the expected loss for a random bet on 'over' was more than twice the size of the expected loss of a bet on 'under'. They suggested that punters exhibited a behavioural bias, as they appeared to show a preference for betting on 'over'. Oikonomidis and Johnson (2008) analysed a similar data set and identified a similar over/under bias. However, this bias was shown to decrease through time. This was shown not to arise from changes in goal-scoring frequency, but from bookmakers offering significantly lower odds on 'under' (and higher on 'over') in later years.

3.6 Subjective estimations

Various public media, including newspapers, radio and television programmes and web sources provide punters with predictions regarding football events. Even though the mechanisms by which such advice is transmitted may be different, they all reflect subjective estimations of a person or a group of persons, involved with the world of football. A number of studies have challenged the value of these subjective predictions. Their findings are presented below.

Using a sample of 1,694 English football games, Forrest and Simmons (2000) tested the value of newspaper tipsters' services. Even though some forecasting ability was observed (tipsters' predictions were significantly better than random), Forrest and Simmons concluded that tipsters fail to adequately account for publicly available information concerning teams' strengths. Moreover, they tested whether tipsters' predictions were informative after performance measures of team strength were considered and found this to be the case for only one of the tipsters. Thus, these studies offer no convincing evidence that the guidance offered by tipsters is of great value, unless there is no other information available. Andersson *et al.* (2005) organized a survey concerning football predictions for the 2002 World Cup. Taking part were 251 participants, varying from football fans, journalists and coaches to non-experts. Both experts and non-experts were found to predict better than random, but there was no evidence that experts predicted more accurately than non-experts. Surprisingly, a simple prediction rule, based on world rankings, achieved superior predictions than most of the participants.

Spann and Skiera (2009) explored data that included stock prices on a prediction market, sports journal tipster predictions and bookmaker odds regarding football matches played in the German Bundesleague in the period 1999–2002. They found that predictions based on betting odds and the prediction market achieve approximately the same level of accuracy (which is significantly more accurate than tipsters' forecasts). However, some suggested rules for combining these three prediction sources led to improved forecasts which could be profitable in a 'friendlier' jurisdiction, where transaction costs are lower than the 25 per cent faced in the German market. However, more research is required to confirm this result, as the sample of bets employed in the study was limited.

3.7 Betting in-running

Betting during the course of a football game has become extremely popular; Hill (2009) estimated that half of betting activity takes place in-running. As a result, a number of researchers have examined whether the market is able to efficiently incorporate the continuous, dynamic flow of information arising from live football action.

Using a data set of 4,000 English football matches, Dixon and Robinson (1998) developed a birth-process model to predict football outcomes during the course of a football game, based on the home advantage, the attacking and defensive abilities of the teams, the current score and the time remaining until the end of the game. They found that the scoring rate increased through the game and, therefore, that a non-homogeneous process is appropriate to model the expected result. They also found that the scoring rate depended on the current score, and, in particular, the scoring rates of both teams decreased significantly when the home side held a narrow lead. Dixon and Robinson (1998) tested the model against spread-betting odds and found some evidence of inefficient pricing.

In a more recent study, Croxson and Reade (2008) used in-running data (concerning 1,206 football matches played in various competitions) from the betting exchange market. The response of the market to significant updates of information (i.e. goals scored) was compared to updated theoretical odds based on a Poisson model fitted to historical data. No evidence of inefficiency in price setting was found, while no relationship between liquidity and inefficient pricing was identified.

3.8 Summary

As indicated above, odds have been shown to be successful predictors of football outcomes. In this section, it is asserted that their forecasting power is comparable to that of sophisticated, fundamental models that utilize a range of publicly available information. However, some researchers find evidence concerning the existence of semi-strong inefficiency and others do not, both sets employing similar information across different samples. In parallel, odds have been shown to be more efficient in responding to certain types of information. It seems to be clear that in recent years odds-setting has improved, posing a more difficult challenge to those intent on making a profit from betting on football.

4 Strong-form efficiency

A market exhibits strong-form information efficiency if prices fully reflect all publicly and privately held information (Fama, 1970). Consequently, the football betting market is strong-form inefficient if some market operators possess 'superior' information regarding the 'true' odds of football events. Thus, in football, where allegations of match fixing have surfaced in recent years, it may be the case that individuals involved in match fixing are exclusively aware of the fact that some outcome is very likely to occur; they may, therefore, use this information to make a profit from betting – leading to a strong-form inefficient market. In the following section we examine cases of match fixing in football and their association with the betting market.

Numerous betting-related football scandals have been revealed through the years. In early 1960s, a group of players, organized by Jimmy Gauld, fixed the outcomes of several football games in order to profit from betting against bookmakers (Preston and Szymanski, 2003). More recently, the goalkeepers of Liverpool and Wimbledon were accused of accepting bribes to fix games between 1993 and 1994, which was linked with Asian betting syndicates (Preston and Szymanski, 2003). In Italy, individuals have confessed to attempting to fix games, with the purpose of profiting from betting, during 1979/1980 period, and in Germany, the referee Robert Hoyzer was convicted of match rigging in 2005. Betting-related match fixing cases have also been identified in Malaysia (Hill, 2009).

The consequences of match fixing are detrimental to all participants in the betting market, excluding the match fixers themselves. From the bookmakers' point of view, match fixing may lead to significant betting losses (to the match fixers) and the demand for betting from honest punters may decrease if such events trigger doubts concerning the fairness of the game (Hosmer-Henner, 2010). From the punters' perspective, they may lose money directly to the match fixers in the betting exchange market. Thus, football fans and authorities, punters and the betting industry all have an interest in keeping the game 'clean'. Consequently, we discuss below how match fixing can still take place.

4.1 How are matches fixed?

Hill (2009) conducted over 220 interviews with match fixers, players, referees, sports and law enforcement officials and agents in the gambling industry. He created the 'Fixed Match Database', which includes matches presumed to have been played honestly together with 130 legally certified fixed football matches. The database also includes a sample of 117 players who were approached to fix matches (of whom 24 refused). Hill (2009) shows that for a match fixer to be successful, five stages of corruption have to be completed successfully: 'access', 'set up', 'calling the fix', 'performance' and 'payment'. Initially, the match fixer needs to gain access (directly or through an agent) to at least one influential player, who will then organize a network within the team to undertake the match fixing operation. Then, the most suitable way to set up the arrangement has to be identified (who to approach and how). Hill (2009) shows that, depending on the type of game, the match fixer's approach may be more or less personal. Hill's data suggests that corrupted players tend to underperform to achieve the desired outcome and that match fixers pay some money to the players in advance, but the main payment is made after the desired result is achieved (usually in cash and not through any sophisticated network).

Match fixing by gambling syndicates has been documented, even without the involvement of the main participants in the game (i.e. players, coaches, managers and referees). For example, in 1999 an Asian gambling syndicate sabotaged the lighting systems of English football stadiums while the score of the game was favourable to them, resulting in high gambling profits (Hosmer-Henner, 2010).

To profit from betting on a fixed game, the match fixers also need to operate successfully in the betting market. They need to explore the type of bet that will maximize their profit, identify ways to place high stakes (usually disguising their identity) and to ensure that the players will bring about exactly the desired outcome. According to Hill (2009), in order to remain unnoticed, the match fixers choose games in which the betting market liquidity is high, so their actions will not cause significant moves in the odds. Alternatively, they may bet on favourites and profit by ensuring the realization of an expected result, which will naturally cause little suspicion. Spreading rumours concerning fixing related to the team opposing the one actually approached has also been documented; the aim being to stimulate bets on the opposing team and increase the odds on the team they intend to bet. Finally, the results of Hill's (2009) research suggests that the corruptors are also more likely to enter the market late, so as not to signal information that may decrease the odds of the team they intend to back.

4.2 Identifying potentially fixed games

The Union of European Football Associations (UEFA) has set up betting fraud detection systems across Europe in order to investigate 27,000 matches played across all the associations (see http://news.bbc.co.uk/sport2/hi/football/europe/7964790.stm). Betting companies and have also established 'early warning systems' aimed at the identification of fixed matches (Hill, 2009). It might be thought that the analysis of fundamental statistics regarding football events may also assist in the discovery of fixed games. However, Hill (2010) compared 137 fixed matches to 120 matches that were (or at least assumed to have been) played honestly and uncovered little statistical evidence against dishonest players. The problem is that players intending to fix a result appear to prefer to under-perform rather than to conduct serious, notable errors, such as own goals or conceding penalties (which are too readily identified). Nevertheless, it was found that the goal-scoring rate in fixed games was higher at the beginning of the game and decreased near its end; the opposite trend to that observed in non-fixed games (Dixon and Robinson, 1998).

4.3 Summary

There is evidence that fixing football matches is possible, and that it has taken place in different types of competition in many countries. The large size of the gambling market induces match fixers to attempt to profit from fixing football outcomes. In more recent years, football and betting authorities have established intelligence systems to identify suspicious games, meaning the match fixers' task has possibly become more difficult. However, it is very difficult for authorities to prevent it completely, as long as the betting market is characterized by high liquidity. Thus, it seems likely that strong-form inefficiency in the football betting market will continue to exist.

5 Conclusions

The study of the literature confirms that the football betting market, like most other markets, exhibits several types of information inefficiency. Thus, information concerning the odds, fundamental data associated with teams' performances, psychological biases and even inside information may be utilized by gamblers to achieve positive returns. However, it is clear that the football betting market is dynamic, and the observed inefficiencies are not necessarily persistent through time. Moreover, it is likely that any strategy that aims to exploit market inefficiencies will be subject to difficulties associated with implementation. Thus, even though opportunities for profit theoretically exist, only the fastest, most efficient and highly determined players are likely to convert theory to practice and benefit from inefficient pricing in the football betting market.

References

Andersson, P., Edman, J. and Ekman, M. (2005), Predicting the World Cup 2002 in soccer: performance and confidence of experts and non-experts, *International Journal of Forecasting* 21, 565–576.

Baxter, M. and Stevenson, R. (1988), Discriminating between the Poisson and the negative binomial distributions: an application to goal scoring in association football, *Journal of Applied Statistics* 15, 347–348.

Bernile, G. and Lyandres, E. (2008), Understanding investor sentiment: the case of soccer, Working Paper.

Cain, M., Law, D. and Peel, D. (2000), The favourite–longshot bias and market efficiency in UK football betting, *Scottish Journal of Political Economy* 47, 25–36.

Clarke, S.R. and Norman, J.M. (1995), Home ground advantage of individual clubs in English soccer, *Journal of the Royal Statistical Society, Series D* 44 (4), 509–521.

Courneya, K.S. and Carron, A.V. (1992), The home advantage in sport competitions: a literature review, *Journal of Sport and Exercise Psychology* 14, 13–27.

Croxson, K. and Reade, J. (2008), Information and efficiency: goal arrival in soccer betting, Working paper, Oxford University.

DCMS (2001), *Gambling Review Report*, 'The Budd Committee Report', available at www.culture.gov.uk/global/publications/archive_2001/gamb_rev_report.htm.

Deschamps, B. (2008), Betting markets efficiency: evidence from European football, *The Journal of Gambling Business and Economics* 2 (1), 66–76.

Deschamps, B. and Gergaud, O. (2007), Efficiency in betting markets: evidence from English football, *The Journal of Prediction Markets* 1, 61–73.

Dixon, M.J. and Coles, S.G. (1997), Modelling Association Football scores and inefficiencies in the football betting market, *Applied Statistics* 46, 265–280.

Dixon, M. and Pope, P. (2004), The value of statistical forecasts in the UK association football betting market, *International Journal of Forecasting* 20, 697–711.

Dixon, M.J. and Robinson, M.E. (1998), A birth process model for association football matches, *The Statistician* 47, 523–538.

Fama, E.F. (1970), Efficient capital markets, a review of theory and empirical work, *The Journal of Finance* 25(2), 383–417.

Forrest, D. and Simmons, R. (2000), Forecasting sport: the behaviour and performance of football tipsters, *International Journal of Forecasting* 16, 317–331.

Forrest, D. and Simmons, R. (2002), Outcome uncertainty and attendance demand in sport: the case of English soccer, *The Statistician* 51, 229–241.

Forrest, D. and Simmons, R. (2003), Sport and gambling, *Oxford Review of Economic Policy* 19 (4), 598–611.

Forrest, D. and Simmons, R. (2008), Sentiment in the betting market on Spanish football, *Applied Economics* 40, 119–126.

Forrest, D., Goddard, J. and Simmons, R. (2005), Odds-setters as forecasters: the case of English football, *International Journal of Forecasting* 21, 551–564.

Franck, E., Verbeek, E. and Nuesch, S. (2009), Inter-market arbitrage in sports betting, NCER Working Paper, No. 48, Series from the National Centre for Econometric Research.

Franck, E., Verbeek, E. and Nuesch, S. (2010), Prediction accuracy of different market structures: bookmakers versus a betting exchange, *Sports Forecasting* 26 (3), 448–459.

Garicano, L., Palacios, I. and Prendergast, C. (2005), Favoritism under social pressure, *Review of Economics and Statistics*, 87, 208–216.

Global Betting and Gaming Consultants (2001), *1st Annual Review of the Global Betting and Gaming Market*, West Bromwich, Global Betting and Gaming Consultants.

Goddard, J. (2005), Regression models for forecasting goals and match results in association football, *International Journal of Forecasting* 21, 331–340.

Goddard, J. and Asimakopoulos, I. (2004), Forecasting football results and the efficiency of fixed-odds betting, *Journal of Forecasting* 23, 51–66.

Goddard, J. and Thomas, S. (2006), The efficiency of the UK fixed-odds betting market for Euro 2004, *International Journal of Sport Finance* 1, 21–32.

Graham, I. and Stott, H. (2008), Predicting bookmaker odds and efficiency for UK football, *Applied Economics* 40, 99–109.

Greenhough, J., Birch, P.C., Chapman, S.C. and Rowlands, G. (2002), Football goal distributions and extremal statistics, *Physica A* 316, 615–624.

Hill, D. (2009), How gambling corruptors fix football matches, *European Sport Management Quarterly* 9 (4), 411–432.

Hill, D. (2010), Strategies of corruption: how corrupt football players and referees perform fixed matches. In M. Kreutner (ed.), *Practice Meets Science: Contemporary Anti-Corruption Dialogue IACSS 2009* (35–50). Manzsche Verlags- und Universitatsbuchhandlung.

Hosmer-Henner, A. (2010), Preventing game fixing: sports books as information markets, *Gaming Law Review and Economics* 14 (1), 31–38.

Jones, P., Clarke-Hill, C.M. and Hillier, D. (2000), Viewpoint: back street to side street to high street to e-street: sporting betting on the Internet, *International Journal of Retail & Distribution Management* 28 (6), 222–227.

Jones, P., Hillier, D., Turner, D. and Comfort, D. (2004), Betting on the exchanges: changing customer relationships in the sports betting market in the UK, *Management Research* 27, 95–103.

Karlis, D. and Ntzoufras, I. (2003), Analysis of sports data by using bivariate Poisson models, *The Statistician* 52 (3), 381–393.

Karlis, D. and Ntzoufras, I. (2009), Bayesian modelling of football outcomes: using the Skellam's distribution for the goal difference, *IMA Journal of Management Mathematics* 20, 133–145.

Kuypers, T. (2000), Information and efficiency: an empirical study of a fixed odds betting market, *Applied Economics* 32, 1353–1363.

Leitner, C., Zeileis, A. and Hornik, K. (2008), Who is going to win the Euro 2008? (A statistical investigation of bookmakers odds). *Report 76*, Department of Statistics and Mathematics, Wirtschaftsuniversität Wien, Research Report Series.

Leitner, C., Zeileis, A. and Hornik, K. (2009), Forecasting the winner of the UEFA Champions League 2008/2009. In R. Koning and P. Scarf (eds), *Proceedings of the 2nd International Conference on Mathematics in Sport – IMA Sport 2009* (pp. 94–99), The Institute of Mathematics and its Applications.

Luckner, S. and Weinhardt, C. (2008), Arbitrage opportunities and market-making traders in prediction markets, *IEEE Joint Conference on E-Commerce Technology (CEC '08) and Enterprise Computing, E-Commerce and Eservices (EEE '08)*, 21–24 July, Washington, DC.

Levitt, S.D. (2004), Why are gambling markets organised so differently from financial markets?, *Economic Journal* 114, 223–246.

Maher, M. (1982) Modelling association football scores, *Statistica Neerlandica* 36, 109–118.

Malaric, R., Katic, T. and Sabolic, D. (2008), The market efficiency of the soccer fixed odds internet betting market, *Applied Economics Letters* 15 (3), 171–174.

Marshall, B.R. (2009), How quickly is temporary market inefficiency removed?, *The Quarterly Review of Economics and Finance* 49 (3), 917–930.

Mintel (2001), *Online Betting, Mintel Intelligence Report*, London, Mintel International Group.

National Gambling Impact Study Commission (1999). Final Report, Washington, DC.

Oikonomidis, A. and Johnson, J.E.V. (2008), Exploring the efficient market hypothesis through time: evidence from the total goals football betting market, *Symposium of the Southern Economic Journal: Gambling, Prediction Markets and Policy Implications*, September, Nottingham.

Page, L. (2009), Is there an optimistic bias on betting markets?, *Economics Letters* 102 (2), 70–72.

Paton, D. and Vaughan Williams, L. (1998), Do betting costs explain betting biases?, *Applied Economics Letters* 5, 333–335.

Paton, D. and Vaughan Williams, L. (2005), Forecasting outcomes in spread betting markets: can bettors use 'quarbs' to beat the book?, *Journal of Forecasting* 24, 139–154.

Paton, D., Siegel, D.S. and Vaughan Williams, L. (2002), A policy response to the e-commerce revolution: the case of betting taxation in the UK, *Economic Journal*, 112 (480), F296–F314.

Pope, P. and Peel, D.A. (1988), Information, prices and efficiency in a fixed odds betting market, *Economica* 56, 323–341.

Preston, I. and Szymanski, S. (2003), Cheating in contests, *Oxford Review of Economic Policy* 19, 612–624.

Reep, C., Pollard, R. and Benjamin, B. (1971), Skill and chance in ball games, *Journal of the Royal Statistical Society, Series A* 131, 581–585.

Rodney, P.J. and Weinbach, A.P. (2009), Are Behavioral biases consistent across the Atlantic? The over/under market for European soccer, *The Journal of Gambling Business and Economics* 3 (2), 89–101.

Sauer, R. (1998), The economics of wagering markets, *Journal of Economic Literature* 36, 2021–2064.

So, H.T. and Kwok, N.K. (2007), No longer an entertainment: social costs of problem gambling in Hong Kong, *Journal of Macau Studies* 41, 94–104.

Spann, M. and Skiera, B. (2009), Sports forecasting: a comparison of the forecast accuracy of prediction markets, betting odds and tipsters, *Journal of Forecasting* 28, 55–72.

SportBusiness Group (2001), *Online Gambling – the Regulation of Global Sports Betting: Executive Summary*, available at www.sportbusiness.com.

Strumbelj, E. and Sikonja, M.R. (2010), Online bookmakers' odds as forecasts: the case of European soccer leagues, *Sports Forecasting* 26 (3), 482–488.

Strumbelj, E., Sikonja, M.R. and Kononenko, I. (2009), Learning betting tips from users' bet selections. In P. Perner (ed.), *Machine Learning and Data Mining in Pattern Recognition* (pp. 678–688), Springer.

Vaughan Williams, L. (2005), *Information Efficiency in Financial and Betting Markets*, Cambridge University Press.

Vlastakis, N., Dotsis, G. and Markellos, R.N. (2008), Nonlinear modelling of European football scores using support vector machines, *Applied Economics* 40 (1), 111–118.

Vlastakis, N., Dotsis, G. and Markellos, R.N. (2009), How efficient is the European football betting market? Evidence from arbitrage and trading strategies, *Journal of Forecasting* 28, 426–444.

Wong, Y.-L.R., Leung, Y.K.T. and Lau, C.W.D. (2009), Behind the allure of gambling: a qualitative exploration of the existential yearnings of Chinese men with problem gambling in Hong Kong, *International Gambling Studies* 9 (3), 189–205.

Woodland, L.M. and Woodland, B.M. (1994), Market efficiency and the favorite–long shot bias: the baseball betting market, *Journal of Finance* 49 (1), 269–280.

16 The prediction market for the Australian Football League

Adi Schnytzer

1 Introduction

The legal Australian Rules football prediction market is less than two decades old, even though punters have doubtless been betting among themselves on their preferred teams for more than a century. As the legal market is young, there is little academic literature analyzing the market. The purpose of this chapter is to make a contribution to this literature, both by subjecting new empirical models to scrutiny and at the level of methodology.

Taking advantage of a novel micro-level data set that includes detailed per-game player statistics, predictions are presented and tested out-of-sample for the simplest kind of bet: fixed odds win betting. It is shown that player-level statistics may be used to yield very modest profits net of transaction costs over a number of seasons, provided some more global variables are added to the model. In particular, the numbers of kicks, marks, handballs and so on obtained by players in a game does not give sufficient information to provide profits in a simple framework, but adding a variable indicating that a team has an a priori home-ground advantage in the game is sufficient to generate profits. A comparison of different specifications of the linear probability model (LPM) versus conditional logit (CLOGIT) regressions reveals that the LPM generally outperforms CLOGIT in terms of profitability.

The methodological question posed here is somewhat obvious: is it necessarily the case that a better regression, in terms of such criteria as adjusted R^2 or log likelihood and statistical significance of explanatory variables, will always lead to increased profitability when the predictions are used to bet in the market? The results presented in this chapter refute this somewhat appealing hypothesis. It is shown that adding a variable that measures a team's performance prior to the current game in the relevant season, while unambiguously improving the regression – be it LPM or CLOGIT – reduces profits (or increases losses) in 9 out of 14 annual cases and turns overall profits into losses.

That this point is not entirely obvious may be understood from a careful reading of the pioneering paper on the Australian Football League (AFL) prediction market (Bailey and Clarke, 2004). Thus, in discussing criteria for the inclusion of variables in their model, they write:

> Variables included in the multiple regression were home ground advantage, interstate travel, ground familiarisation, team quality and current form, with all variables being statistically significant with a *p*-value <0.0001. We have found that using such a stringent significance level creates more robust predictors.

What makes this paper so interesting is that it presents a very thorough analysis of the complexities involved in making predictions sufficiently accurate to permit profitable betting in the AFL prediction market. Discussed are the various explanatory variables and optimization of their measurement. For example, the decomposition of home-ground advantage into home-player familiarity with the ground, visiting team fatigue in traveling interstate and other factors evidently adds to profits. There is also testing of the optimal way to predict explanatory variables such as past performance: does one use moving averages of exponential smoothing or some other technique? The problem from an academic viewpoint is that the paper suppresses actual coefficient values and other details for commercial reasons.

The subsequent literature is far more specialist in nature, and a very brief description of three papers will suffice.[1] Grant and Johnstone (2010) predict game outcomes and simulate betting by pooling forecasts of winning probabilities derived from a web-based football "tipping" competition, which has been conducted by the computer science faculty at Monash University in Melbourne since 1995. They present exhaustive tests of different pooling and betting methods and show that statistically significant, although not large, profits may occasionally be made using this approach, although in the long-term average losses prevail.

Ryall and Bedford (2010), on the other hand, claim that long-run profits are available in this market if a ratings-based forecasting model is adopted. The model used is that of Elo (1978), originally designed for ranking chess players. Over the 2001–2008 AFL seasons they generate a return of investment of 8.8 percent, betting a constant amount on each game and 10 percent using a Kelly system. These returns are greater than those presented here, but the method adopted is highly computer intensive and may be impractical if rankings are to be updated after each round.[2] If this model is indeed successful, it would presumably yield even better results if rankings were regularly updated.

Sargent and Bedford (2010) show how exponentially smoothed, one-step forecasts of AFL player performance data are improved by first applying a non-linear smoother to the raw data. In this respect, their paper builds upon Bailey and Clarke (2004) in its analysis of exponential smoothing as yielding improved forecasts over simple and moving averages. Player performance is defined as an index based upon several player-level statistics of the kind used in the paper (kicks, handballs, etc.), but no use of the predictions in simulated (or real) betting in the AFL prediction market is presented.

The central feature of the analysis presented here is its attempt at simplicity, if not naivety. Thus, the regressions run are of the simplest kind and the variables used are extremely basic: no attempt is made to index player performance, the emphasis being on the raw data. Further, home-ground advantage is represented by a dummy variable, thus precluding any degrees of advantage. Finally, in order to predict player performance, reliance is upon simple means alone. The reason for this approach is two-fold. First, it is interesting to ask whether profits are obtainable, however modest they may be, without resorting to complications – and the answer turns out to be positive. Second, testing the methodological hypothesis that the better the regression, the more profitable will be the predictions it yields, requires that as many confounding factors as possible be removed from the analysis.

2 Australian Rules football[3]

Australian Rules football, also known as "Australian football," "Aussie rules," "football" or "footy," is a code of football played with a prolate spheroid ball on large oval-shaped fields, with four posts at each end: two tall posts in the centre – "goal posts" – and two shorter outer posts – "behind posts" or "point posts." The playing field may be 135–185 meters long and 110–155 meters wide.

Footy, as it is generally known today, originated in Melbourne in 1858 and was devised to keep cricketers fit during the winter months. The football season runs from March to August (early autumn to late winter in Australia), with finals in September. Some claim that *Marngrook* – a traditional Aboriginal ball game played for millennia in what is now western Victoria – provided the first law-makers of football with some of the fundamentals of Australian Rules football. However, opinion across the footy-loving Australian public is divided as to Marngrook's contribution to the modern game.[4]

A football game consists of four 20-minute quarters, plus time added for stoppages. Most quarters effectively last 25–30 minutes. Each team has 18 players on the field at any given time and four substitutes are available for unrestricted, repeated substitutions as deemed fit by the coach. Since footy allows players to handle the ball as well as kick it, and since there are no off-side rules, the game is in many ways similar to basketball in the speed and extent of scoring. For the 1998–2007 seasons, the average game score per team was 95 points, with a minimum of 23, a maximum of 222 and a standard deviation of 28. A comprehensive introduction to the game is provided at www.footy.com.au/dags/FAQ1v1–5.html and http://en.wikipedia.org/wiki/Australian_rules_football. The official website of the AFL is www.afl.com.au. For videos of game highlights see www.youtube.com/watch?v=xIOvSv9Q1Gk.

3 AFL prediction markets

The three major types of betting market in the AFL are fixed-odds betting for the win, line betting and even-money line betting. While, for the purposes of this chapter, win betting is the focus of attention, a brief description of the associated markets is provided in this section.[5]

A typical line wager in the AFL requires that the bettor risk $1 for the chance to receive around $1.9.[6] This $1.9-for-$1 dividend requires that bettors pick winners in 52.63 percent of bets to break even.[7] In the event the outcome is identical to the line, known as a "push" or a "no bet," the gambler's wager is refunded.

The even-money line (or *points*) wager is quite similar to the line wager, yet the dividend is always $2. Therefore, in this case, the percentage of winning bets required to break even is 50 percent. Different bookmakers offer different point spreads on the AFL. These spreads are between six and ten points, i.e., spreads of three and five points on either side of the line. The result of a match falls in the six-points spread around 6 percent of the games; hence, by offering a six-points spread at even-money, the bookmaker retains around 6 percent of his turnover, which is around 0.5 percent higher in the six-points spread than its equivalent in the line market,[8] and around 2 percent higher than in the fixed-odds market.

The even-money line bet is based on a bid/ask spread, which is the difference between the price available for an immediate sale (bid) and an immediate purchase (ask). For example, if the even-money line is 35 for team A to win and 28 for team B to lose, the bettor can either bet on team A to win by 35 or more, or on team B to lose by 28 or less. No one can bet on the spread between 29 and 34 points, the range of possible bets in which the bookmaker wins all bets. This is parallel to a bid of 28 and an ask of 35, where the broker makes his money.

The fixed-odds win wager in the AFL, which is the subject of simulation in this chapter, requires the bettor to risk $1 for the chance to receive a fixed sum if successful.[9] As in the above prediction markets, the bookmaker sets odds to earn around 5 percent of the total bet if his book is balanced.[10] Nevertheless, unlike the line and even-money line betting markets, there is no certain percentage of winning bets necessary to break even in the fixed-odds market, since the range of actual payouts is huge.[11] This market provides the central focus of this chapter.[12]

Other betting methods are also available in the AFL: draw, where the bettor bets on the chance that the final result will be a draw; point spread in ten-point gaps; 1–39 and 40+, where the bettor bets on the chance that the point spread will be between 1 and 39 points or 40 and above; highest scoring quarter; first goal scorer in each quarter; most goals kicked; most free kicks; and also various future odds bets, including different medals, Premiership, final eight, highest placed Victorians or non-Victorians, team to reach the Grand Final, first coach to depart and many others.[13]

4 Data and analysis

The raw data used in this chapter are derived from publicly available sources, i.e., internet-based sports statistical information. The game data come from the official league website (www.afl.com.au) and from http://stats.rleague.com, while betting data are from www.sportsbetting.com.au.[14] The data consist of individual player statistics for all AFL games from the first round of the 1998 season through the Grand Final of 2007, team performances, dates, grounds and the last available fixed odds for the win for each team. This amounts to 81,400 player-level observations over 1,850 games.

The variables employed in the prediction models are defined as:[15]

win$_{jk}$ = 1 if team *j* won game *k* and 0 if it lost or (very rarely[16]) drew. This is the dependent variable in all regressions
kicks$_{ijk}$ = the number of kicks obtained by player *i* of team *j* in game *k*
marks$_{ijk}$ = the number of marks taken by player *i* of team *j* in game *k*
handballs$_{ijk}$ = the number of handballs provided by player *i* of team *j* in game *k*
tackles$_{ijk}$ = the number of tackles by player *i* of team *j* in game *k*
clangers$_{ijk}$ = the number of clangers for which player *i* of team *j* was responsible in game *k*
rebound50s$_{ijk}$ = the number of times player *i* of team *j* retrieved the ball and sent it out of the opposing team's 50 meter attacking zone in game *k*
hitouts$_{ijk}$ = the number of hitouts obtained by player *i* of team *j* in game *k*
clearances$_{ijk}$ = the number of times player *i* of team *j* cleared the ball out of defense in game *k*
freesfor$_{ijk}$ = the number of free kicks received by player *i* of team *j* in game *k*
freesagainst$_{ijk}$ = the number of free kicks given away by player *i* of team *j* in game *k*
dummy_home$_{jk}$ = 1 if team *j* has an a priori home-ground advantage in game *k*, 0 otherwise
neutral$_{jk}$ = 1 if team *j* is playing game *k* at a neutral ground, 0 otherwise
clinch_1$_{jk}$ = 1 if team *j* has already clinched a place in the finals before the start of game *k*, 0 otherwise
elim_1$_{jk}$ = 1 if team *j* has already been eliminated from the finals before the start of game *k*, 0 otherwise
winpct_1$_{jk}$ = the proportion of games that team *j* has won this season prior to game *k*.[17]

In addition to these data, we have the bookmakers' odds for a win bet on each of the teams playing. The process of econometric prediction and out-of-sample betting simulation is as follows:

1 Four parallel pairs of regression specifications are run, one quartet using the LPM and the other using McFadden's (1973) CLOGIT. The first pair of regressions contain player-level variables only and these are shown for the whole sample in Table 16.1; regressions LMP 1 and CLOGIT 5, respectively. For the purposes of prediction, however, the regressions are run on the data subset containing all observations from the first round of 1998 through the 2000 Grand Final. These regressions are used to predict the winning probabilities of the teams in round 1 of 2001 by substituting the mean values of the player-level explanatory variables for the 1998–2000

Table 16.1 Regression results for the entire sample (1998–2007)

	LPM 1	*LPM 2*	*LPM 3*	*LPM 4*	*CLOGIT 5*	*CLOGIT 6*	*CLOGIT 7*	*CLOGIT 8*
Variables	*Coefficient (t-stat)*	*Coefficient (t-stat)*	*Coefficient (t-stat)*	*Coefficient (t-stat)*	*Coefficient (z-stat)*	*Coefficient (z-stat)*	*Coefficient (z-stat)*	*Coefficient (z-stat)*
Kicks	0.0118	0.0112	0.0108	0.0105	0.0495	0.0483	0.0467	0.0462
	(20.45)	(19.72)	(18.67)	(18.04)	(20.52)	(19.75)	(18.43)	(17.47)
Marks	0.0070	0.0072	0.0071	0.0071	0.0317	0.0335	0.0338	0.0369
	(7.86)	(8.34)	(8.05)	(7.97)	(8.48)	(8.82)	(8.61)	(8.99)
Handballs	0.0048	0.0041	0.0040	0.0037	0.0238	0.0209	0.0201	0.0190
	(7.98)	(6.85)	(6.65)	(6.00)	(9.23)	(8.02)	(7.41)	(6.70)
Tackles	0.0043	0.0044	0.0041	0.0043	0.0154	0.0157	0.0159	0.0170
	(4.03)	(4.16)	(4.78)	(3.94)	(3.35)	(3.38)	(3.28)	(3.38)
Clangers	–0.0276	–0.0278	–0.0270	–0.0261	–0.1167	–0.1212	–0.1168	–0.1151
	(–19.92)	(–20.45)	(–19.42)	(–18.53)	(–19.35)	(–19.80)	(–18.39)	(–17.31)
Rebound50s	–0.0141	–0.0128	–0.0121	–0.116	–0.0663	–0.0622	–0.0590	–0.0580
	(–13.66)	(–12.61)	(–11.64)	(–11.11)	(–14.95)	(–13.83)	(–12.65)	–(11.91)
Hitouts	0.0027	0.0025	0.0024	0.0023	0.01099	0.0102	0.0098	0.0096
	(6.68)	(6.11)	(5.86)	(5.64)	(6.60)	(6.04)	(5.57)	(5.24)
Clearances	0–.0086	–0.0078	–0.0075	–0.0074	–0.04134	–0.0389	–0.0380	–0.0388
	(–7.08)	(–6.54)	(–6.16)	(–6.02)	(–8.22)	(–7.61)	(–7.17)	(–7.01)
Freesfor	–0.0123	–0.0140	–0.0144	–0.0142	–0.0502	–0.0589	–0.0593	–0.0590
	(–6.33)	(–7.36)	(–7.38)	(–7.19)	(–6.27)	(–7.25)	(–7.04)	(–6.69)
Freesagainst	0.0121	0.0150	0.0145	0.0139	0.0506	0.0651	0.0638	0.0629
	(5.79)	(7.31)	(6.93)	(6.56)	(5.79)	(7.33)	(6.93)	(6.52)
Dummy_home		0.2081	0.2049	0.2102		0.8467	0.8586	0.9272
		(49.43)	(47.68)	(48.48)		(47.72)	(46.61)	(48.02)
Neutral		0.1052	0.1098	0.1147				
		(25.12)	(25.68)	(26.53)				

Clinch_1			0.1665	0.1006			1.2217	0.7590
			(19.19)	(11.47)			(17.28)	(10.57)
Elim_1			−0.1700	−0.0960			−1.7097	−0.9258
			(−23.04)	(−12.69)			(−27.01)	(−14.10)
Winpct_1				0.3006				2.5817
				(39.41)				(51.38)
Adj or pseudo R^2	0.0180	0.0477	0.0581	0.0821	0.0147	0.0371	0.0530	0.0852
No. observations	81,400	81,400	77,440	73,918	80,740	80,740	76,780	7,326

Note

All coefficients are significantly different from zero at better than 0.1 percent. The number of observations differs because winpct_1 is undefined for round 1 of each season and because CLOGIT drops the 15 drawn games from the regression. In no LPM regressions were estimated variances negative, all predicted winning probabilities lying in [0,1]. Thus, some observations are lost where estimated variances are zero.

period into the obtained regression results. The second pair of regressions add two dummy variables, the first indicating whether or not the home team has an a priori home-ground advantage, and the second indicating whether or not the stadium in which the current game is being played is a neutral ground, offering no a priori advantage to either side.[18] These are regressions LPM 2 and CLOGIT 6, shown for the whole data set in Table 16.1. Note that "neutral" does not appear in any CLOGIT regression because it must always, by definition, receive the same value for both teams in a game and the conditional logit regression conducts its estimation by distinguishing between the two teams in a game exclusively. Regressions LPM 3 and CLOGIT 7 add to the extant explanatory variables two team-level dummy variables which indicate whether or not the team has clinched a place in the finals or whether the team has definitely been eliminated from the finals race immediately prior to the game to be played, respectively. Finally, regressions LPM 4 and CLOGIT 8 add a further team-level variable which measures the proportion of wins accumulated by the team so far in the current season prior to the current game.

2 On the basis of these regressions, predicted winning probabilities for the teams in each game of round 1 of the 2001 season are calculated as follows. For each player in the team, each regression predicts a probability which may be interpreted as that player's predicted contribution to the team's winning probability. In the case of the CLOGIT regressions, these probabilities sum to 1 for each game. Thus, summing them across players in any given team yields the predicted winning probability for that team. The linear probability model requires an extra step since probabilities do not generally sum to 1 for each game.[19] Accordingly, these predictions are normalized over each game and the resultant sums per team taken as the predicted winning probabilities for the relevant team.

3 Given the teams' predicted winning probabilities and the bookmakers' prices for a win bet on each team, the simulated betting is on those teams for which the predicted winning probability exceeds 0.5 (i.e., the predicted favorites in the game) and the amount bet is in proportion to the predicted winning probability. This betting system is adopted as it is the method adopted by many Australian professional punters.[20]

4 The results for round 1 of 2001 being now known, as it were, the data for this round are added to the data set and all the regressions re-run to predict the winning probabilities for each team in round 2 of 2001, and betting is again simulated. This process continues with new regressions being run round by round until the end of 2007 and the total results of simulated betting calculated year by year. These results are shown in Tables 16.2 and 16.3.[21]

Prior to a discussion of the betting results, some discussion of Table 16.1 is in order. While these specific regression results are for the entire data set and thus do not feature in any of the simulations, they turn out to be representative of

Table 16.2 Betting simulation results for the LPM

Year	*Specification 1*			*Specification 2*			*Specification 3*			*Specification 4*		
	Amount bet	*Profit*	*Cumulative rate of return*	*Amount bet*	*Profit*	*Cumulative rate of return*	*Amount bet*	*Profit*	*Cumulative rate of return*	*Amount bet*	*Profit*	*Cumulative rate of return*
2001	596.96	−36.02	−0.0603	833.34	−60.27	−0.0723	777.58	−43.52	−0.0560	602.37	−11.42	−0.0190
2002	664.05	−257.22	−0.2325	855.07	−143.41	−0.1206	800.24	−61.26	−0.0664	792.97	−53.78	−0.0467
2003	630.81	−189.94	−0.2554	875.29	−42.97	−0.0962	819.84	−10.98	−0.0483	632.32	−25.17	−0.0446
2004	630.41	−66.74	−0.2180	962.47	201.12	−0.0129	859.96	141.82	0.0080	693.72	160.35	0.0257
2005	729.81	−78.50	−0.1932	925.10	109.30	0.0143	815.96	157.16	0.0450	630.39	4.42	0.0222
2006	697.32	15.67	−0.1552	874.42	42.27	0.0199	771.27	56.94	0.0496	651.98	−58.35	0.0040
2007	695.86	128.60	−0.1042	1,048.70	115.63	0.0348	998.55	86.47	0.0559	790.70	−46.59	−0.0064

Notes

All results are out-of-sample. The data set begins in round 1 of 1998 and is updated and the models re-run for each round after round 1 of 2001. The explanatory variables in the four specifications are as follows:

1 kicks, marks, handballs, tackles, clangers, rebound50s, hitouts, clearances, freesfor, freesagainst.

2 kicks, marks, handballs, tackles, clangers, rebound50s, hitouts, clearances, freesfor, freesagainst, dummy_home, neutral.

3 kicks, marks, handballs, tackles, clangers, rebound50s, hitouts, clearances, freesfor, freesagainst, dummy_home, neutral, clinch_1, elim_1.

4 kicks, marks, handballs, tackles, clangers, rebound50s, hitouts, clearances, freesfor, freesagainst, dummy_home, neutral, clinch_1, elim_1, winpct_1.

Table 16.3 Betting simulation results for the CLOGIT model

Year	Specification 5			Specification 6			Specification 7			Specification 8		
	Amount bet	*Profit*	*Cumulative rate of return*	*Amount bet*	*Profit*	*Cumulative rate of return*	*Amount bet*	*Profit*	*Cumulative rate of return*	*Amount bet*	*Profit*	*Cumulative rate of return*
2001	600.33	–19.34	–0.0322	1,140.25	–113.58	–0.0996	1,196.29	–110.77	–0.0926	1,452.08	–132.37	–0.0912
2002	691.04	–236.53	–0.1981	1,227.98	–133.46	–0.1043	1,258.29	–29.10	–0.0570	1,553.14	–30.41	–0.0542
2003	635.70	–154.73	–0.2131	1,218.07	–33.07	–0.0781	1,255.14	3.95	–0.0366	1,455.73	–100.87	–0.0591
2004	623.62	–56.17	–0.1830	1,356.41	285.58	0.0011	1,398.86	200.18	0.0126	1,645.70	173.83	–0.0147
2005	811.18	–21.51	–0.1452	1,348.01	109.39	0.0183	1,349.55	133.48	0.0306	1,616.62	–34.45	–0.0161
2006	702.27	0.84	–0.1199	1,270.09	–54.40	0.0080	1,403.86	–34.61	0.0207	1,646.53	–176.82	–0.0321
2007	743.36	85.11	–0.0837	1,382.04	135.57	0.0219	1,385.57	38.61	0.0218	1,841.62	–50.39	–0.0314

Notes

All results are out-of-sample. The data set begins in round 1 of 1998 and is updated and the models re-run for each round after round 1 of 2001. The explanatory variables in the four specifications are as follows:

5 kicks, marks, handballs, tackles, clangers, rebound50s, hitouts, clearances, freesfor, freesagainst.

6 kicks, marks, handballs, tackles, clangers, rebound50s, hitouts, clearances, freesfor, freesagainst, dummy_home.

7 kicks, marks, handballs, tackles, clangers, rebound50s, hitouts, clearances, freesfor, freesagainst, dummy_home, clinch_1, elim_1.

8 kicks, marks, handballs, tackles, clangers, rebound50s, hitouts, clearances, freesfor, freesagainst, dummy_home, clinch_1, elim_1, winpct_1.

virtually all the other regressions run. Thus, in all regressions run subsequent to the third round of 2002, all player-level variables are statistically significant at better than 1 percent. Prior to that period, the variable measuring the number of free kicks given away by a player in the game is generally statistically insignificant.

Adding variables above the player level consistently improves the regressions. Thus, the variable(s) measuring home-ground advantage are always statistically significant at better than 0.1 percent and always more than double the adjusted (or pseudo) R^2. Adding information regarding whether or not a team has clinched a place in the finals or has been definitely eliminated improves the regressions yet further; the proportion of games won prior to the current game again improves the model.

If the success of predictions is a function of the "goodness" of the econometric model, then it might be expected that models LPM 4 and CLOGIT 8 would perform best, since in every evident statistical respect they appear to be better that their predecessors. This is true not only of the statistical significance of the added variable and adjusted or pseudo R^2 as shown in Table 16.1, but also from regression F-tests and log likelihoods (not shown). Since both sets of regression models are nested, a comparison of these statistics is valid. However, as a perusal of Tables 16.2 and 16.3 indicates, things are not so simple!

The first thing to become clear is that using player-level data alone (at least in the simple way adopted here) in making predictions as a basis for betting yields loses in five of the seven simulated betting seasons, be it using the LPM or the CLOGIT model. Overall, the latter does slightly better with cumulative losses of 8.4 percent as against 10.4 percent. Adding details of the venue to the player-level data improves returns considerably, converting losses into cumulative profits of 3.5 percent for the LPM and 2.2 percent for CLOGIT. Adding details about the teams' progress or otherwise toward a place in the finals adds a further 2 percent to the LPM but has no impact upon the CLOGIT model.

From Tables 16.2 and 16.3 it is clear that the best returns from simulated betting over the entire period derive from either specifications 2 (6) or 3 (7), but definitely not from 4 (8). Thus, it would seem to be that better regressions may not lead to better prediction, per se, although there is a strong reason why this should not be surprising. The generally low percentage of variance in winning probabilities explained by these regressions makes it clear that there is missing variable bias. Further, since the regressions are used for predictions, it is obvious that many relevant variables (e.g., the mental states of the players at game time, their precise physical states, to name just two groups of variables) will never be known. Accordingly, the coefficients in the regressions are inevitably biased as are those in the (real or imputed) regressions of the bookmakers. Now, given the latter, bookmakers' odds are also likely to be biased[22] in many ways, which may only be discovered if a serendipitously "better" prediction model is run. The bettor will thus come up with an edge that seems to defy the formal rule of econometrics. Thus, suppose that in a number of games, the addition of **winpct_1** to, but the absence of say **clinch_1** from, bookmakers' models[23] raises their price *wrongly* above 0.5. Given the large numbers of missing variables in these regressions and the unknown interactions between them, this is certainly not far-fetched. And suppose that a bettor who omits **winpct_1**, but includes **clinch_1**, arrives *correctly*, owing to the imponderable total impact of missing variables bias, at a winning probability of less than 0.5. If there are sufficient cases such as this one, the results shown in specifications 3 and 7, respectively, in Tables 16.2 and 16.3, where returns are generally better than in specifications 4 and 8, which add **winpct_1**, will make perfect sense.

Finally, it may be noted that the results as shown are certainly quite modest, with a best seven-year cumulative return of only 5.6 percent for LPM 3 in Table 16.2. But here, the results would have looked far better had they been framed differently. Instead of training with data from 1998 through 2000 and better from 2001 onwards, suppose that training used data from 1998 through 2003 and began in round 1 of 2004. Then, the cumulative return by the end of 2007 would be a not entirely unrespectable 12.8 percent. Of course, this sleight-of-hand is made possible by the fact that in many models (but noticeably not specifications 4 (8) of Tables 16.2 (16.3)) results improve as time goes by. Whether this is because the predictions improve as more observations are added to the regressions and/or for some other reason(s), is unclear.

5 Conclusions

Taking advantage of a novel micro-level data set which includes detailed per-game player statistics, predictions have been presented and tested out-of-sample for the simplest kind of bet: fixed-odds win betting over the AFL seasons from 2001 through 2007. Data from the beginning of 1998 through the end of 2000 have been used as the source for the initial predictions, while the data have then been updated round-by-round. It has been shown that player-level statistics may be used to yield very modest profits net of transaction costs over this period, provided some more global variables, such as whether or not one team has an *a priori* home-ground advantage and what progress the team has made towards a place in the finals, are added to the model. A comparison of different specifications of the LPM versus CLOGIT regressions reveals that the LPM usually outperforms CLOGIT in terms of profitability. It is further shown that adding significant variables to a regression specification which is clearly superior in econometric terms may reduce the profits derived from the forthcoming predictions.

Notes

1 See Weinberg (2008) for a discussion of other papers.
2 The results cited are based on seasonal updates only.
3 Some of this section is drawn from Weinberg (2008). For further details on the game and its rules, see there.
4 For further details on this dispute, see www.aboriginalfootball.com.au/marngrook.html.

5 The remainder of this section is drawn, with only minor modifications, from Weinberg (2008).
6 In contrast to the US market, the winning dividend per $1 point spread wager in the AFL is not fixed. The range of this dividend in the 2001–2007 period was $1.78–$2.05, while in 70 percent of the games it was $1.9, the mean also being $1.9.
7 The percentage of winning bets (*WP*) necessary to break even, 52.63 percent, is obtained by setting the expected value of the random variable, a gamble $WP \times 0.9 + (1 - WP) \times (-1)$, equal to 0. See, for further discussion, Vergin and Scriabin (1978), Gandar *et al.* (1988), and Dana and Knetter (1994).
8 Response to a query by Hamish Davidson from Sportsbetting.com.au.
9 The range of actual payouts in the 2001–2007 period is $1.02–$14 (mean = $2.39).
10 The average bookmakers' commission in 2001–2006 was 4.5 percent. Bailey and Clarke (2004) noted that the commission could be as low as 2–3 percent.
11 In the very rare event where the outcome is a draw, the fixed-odds bettor wins half the amount he would have won had his team won (see www.bookiering.com). There were only 15 drawn games during seasons 1998 through 2007 inclusive.
12 It should be noted that a prediction market exists also for win betting at fixed odds during the course of each game, but a discussion of this market is beyond the scope of this chapter. See, for example, http://betting.betfair.com/education/sports/04-australian-rules/australian-rules-260908.html for further details.
13 A list of Australian and nearby registered interactive bookmakers and their respective websites can be found at www.betting-ring.com/australia.html.
14 The author wishes to thank Paul Jeffs, who runs http://stats.rleague.com and Hamish Davidson of Sportsbet Pty. Ltd. for providing different subsets of these data in a readily useable form, and Guy Weinberg, Nissim Pinto and Olga Singer for invaluable assistance in organizing the data.
15 Subscripts are used here to facilitate the definitions of the variable, but are dropped thereafter.
16 There are 15 drawn games in the sample.
17 Thus, this variable is never defined for the first round of a season.
18 For a thorough analysis of the subtleties of home and neutral grounds in the AFL, see Schnytzer and Weinberg (2008).
19 It is interesting to note, however, that the predicted probabilities per game always fall between 0 and 1.
20 The author thanks Terry Pattinson (formally Australian sports betting bookmaker and currently Head of In-Play Development for William Hill Plc) for this insight.
21 This method of betting simulation (but with different betting criteria) was used for one season of American football in Zuber *et al.* (1985).
22 See Schnytzer and Weinberg (2008) for evidence of bias in favor of home teams, playing against interstate visiting teams, in states outside of Victoria, but no favorite–longshot bias.
23 Real or implicit.

Bibliography

Bailey, M.J. and Clarke, S.R. (2004). "Deriving profit from Australian Rules football: a statistical approach." In: H. Morton (ed.), *Proceedings of the Seventh Australian Conference on Mathematics and Computers in Sport* (pp. 48–56). Palmerston Nth, New Zealand: Massey University.

Dana, J.D. and Knetter, M.M. (1994). "Learning and efficiency in a gambling market." *Management Science* 40: 1317–1328.

Elo, A.E. (1978). *The Rating of Chessplayers, Past and Present.* Batsford, London.

Gandar, J., Zuber, R., O'Brien, T., and Russo, B. (1988). "Testing rationality in the point-spread betting market." *Journal of Finance* 43: 995–1008.

Grant, A. and Johnstone, D. (2010). "Finding profitable forecast combinations using probability scoring rules." *International Journal of Forecasting* 26: 498–510.

McFadden, D. (1973). "Conditional logit analysis of qualitative choice behaviour." In: P. Zarembka (ed.), *Frontiers in Econometrics*. New York: Academic Press.

Ryall, R. and Bedford, A. (2010). "An optimized ratings-based model for forecasting Australian Rules football." *International Journal of Forecasting* 26: 511–517.

Sargent, J. and Bedford, A. (2010). "Improving Australian Football League player performance forecasts using optimized nonlinear smoothing." *International Journal of Forecasting* 26: 489–497.

Schnytzer, A. and Weinberg, G. (2008). "Testing for home team and favorite biases in the Australian Rules football fixed-odds and point-spread betting markets." *Journal of Sports Economics* 9: 173–190.

Vergin, R. and Scriabin, M. (1978). "Winning strategies for wagering on national football league games." *Management Science* 24: 809–818.

Weinberg, G. (2008). "Efficiency of professional sports wagering markets: the case of Australian Rules football." PhD thesis, Bar Ilan University.

Zuber, R.A., Gandar, J.M. and Bowers, B.D. (1985). "Beating the spread: testing the efficiency of the gambling market for national football league games." *Journal of Political Economy* 93: 800–806.

17 Do experts know more than the crowd?

A case study

Michael A. Smith

1 Introduction

Observed odds distributions and betting patterns in horse racing are often analysed and explained in terms of insider trading activity (Crafts, 1985; Shin, 1991, 1992, 1993). The role of experts who do not have privileged information in shaping or influencing the market has arguably been neglected. This chapter reviews the evidence on the analysis of horse racing fundamentals and the extent to which such analysis can be profitably exploited in horse race betting markets.

Experts in the current context, in the absence of a generally accepted definition, are taken to be the major classes of agent in relation to horse race betting markets, excluding insiders. The discussion therefore centres on the value of publicly available fundamental information in predicting winners; the extent to which such analysis is assimilated in the market; and whether it can be used to execute profitable wagering strategies.

The evidence presented on expert analysis is confined to three groups which use fundamentals: *media forecasters* (comprising newspaper journalists and 'morning line' forecasters,[1] whose selections and odds evaluations respectively are widely published in advance of the races); *private handicappers* (a shorthand term referring to analysts whose race selections are generally not published in the media, a group which includes academics devising models derived *ex post* from historical data, and analysts who make predictions informing their betting activities in real-time, notably professional bettors); and, last but not least, *bookmakers*, whose estimates of winning chances are expressed in the odds they offer to bettors.

Bookmakers are included because, unlike pari-mutuel betting markets, where pre-deduction odds are purely determined by betting demand, bookmakers are market makers who take a view on the chances of race runners when the market opens, although according to Nevison (2009), their estimates are increasingly guided by the prevailing betting exchange odds.[2]

2 Media forecasters

Semi-strong information efficiency as specified by the efficient-markets hypothesis is the degree to which all publicly available information is assimilated into market prices (Fama, 1970). Many studies of stock markets demonstrate rapid assimilation of information by the market (e.g. Krauss and Stoll, 1972; Patell and Wolfson, 1984); others suggest a time lag between publication of information and its assimilation in prices (e.g. Ball and Brown, 1968). A number of studies find that horse race betting markets are efficient in discounting published factual information and published race forecasts based on expert analysis (see Bird and McCrae, 1987 in relation to Australian media experts; Figlewski, 1979 for similar data for the United States; and Vaughan Williams, 2000 regarding UK media selections).

Figlewski (1979) tested the semi-strong informational efficiency of betting markets in respect of 189 races run at Belmont, a US racetrack in the state of New York. He suggested that if the market is efficient, the track odds should incorporate all race-relevant, publicly available information; the information subset used was the published selections for win and position of 14 media forecasters – journalists employed by three major newspapers.

Figlewski utilised a multinomial logit model to derive the log likelihood value of the track odds, and hence their predictive value, and repeated the procedure in respect of track odds *and* forecaster's selections, expressed as 0/1 dummy variables. The latter procedure enabled Figlewski to judge whether media selections added significantly to the information inherent in the odds, by comparing the log likelihood value with that of odds alone.

The null hypothesis therefore tested was that:

$$P(win \mid trackodds) = P(win \mid trackodds + all\ publicity\ held\ information), \tag{17.1}$$

where P is the objective probability of a horse winning. In an efficient market it should not be possible to make more accurate predictions by combining the odds with media selections than can be made using the odds alone. Figlewski found that track odds and media forecasts held substantially more predictive value than no information. This result is to be expected, as favourites alone in any representative sample of horse races will typically yield 25–33 per cent of winners, depending on the average number of runners in the sample, n (which tends to be lower in US racing than in the United Kingdom due to greater restrictions on field size in the USA). The strike rate of the favourite in any representative sample of races is significantly higher than the expected $1/n$ rate associated with random selection.

Figlewski further found that, whereas the log likelihood for media forecasts was lower than that for odds alone, and that for odds and forecasts was lower still, neither figure was significantly lower than the value for odds alone, based on the χ^2 test at the 95 per cent confidence level. Interestingly, however, when Figlewski conducted the tests separately for two distinct groups of bettors, those who bet on-track and those who bet off-track, he found that off-track bettors did not discount media forecasts fully in their betting; knowledge of media information in addition to off-track odds significantly increased the log likelihood value, at the 95 per cent confidence level. Figlewski concluded that the off-track market therefore contained a semi-strong inefficiency. His subsequent out-of-sample predictions based on the multinomial logit coefficients, however, strongly suggested that a tradable strategy based on any discrepancy between track odds and media forecasts could not be profitably exploited, confirming the overall results suggested by log likelihood comparisons.

To be comprehensive, tests of semi-strong efficiency should examine the dynamics of the market in addition to opening or final odds as static structures.

Figure 17.1 shows a stylised version of alternative paths of odds for selections of a hypothetical newspaper result-forecasting column, assumed to be influential in bettors' decision-making, from early morning to race time. Assume that the depicted paths represent the odds of a typical horse (hitherto referred to as a 'tip') selected by the column, and that there are no transaction costs.

M represents media publication time – the time at which tips and initial market prices are simultaneously known to the public (9 a.m. on the morning of race day, for example). *R* is race time, and *A* are the best odds available to bet the tip at time *M*. The lines *ABCD* and *ACD* show alternative paths of odds over time between *M* and *R*.

The conditions for semi-strong efficiency can be breached in two ways. First, if final odds, *F*, do not reflect the true probability of the tip winning. For example, should the fair odds be O_2, a long-term positive return will be earned to wagers struck at *F*, whereas if the fair odds are O_1, this will lead to long-term negative returns. *F*=*D*=fair odds is a necessary but not sufficient condition for semi-strong efficiency.

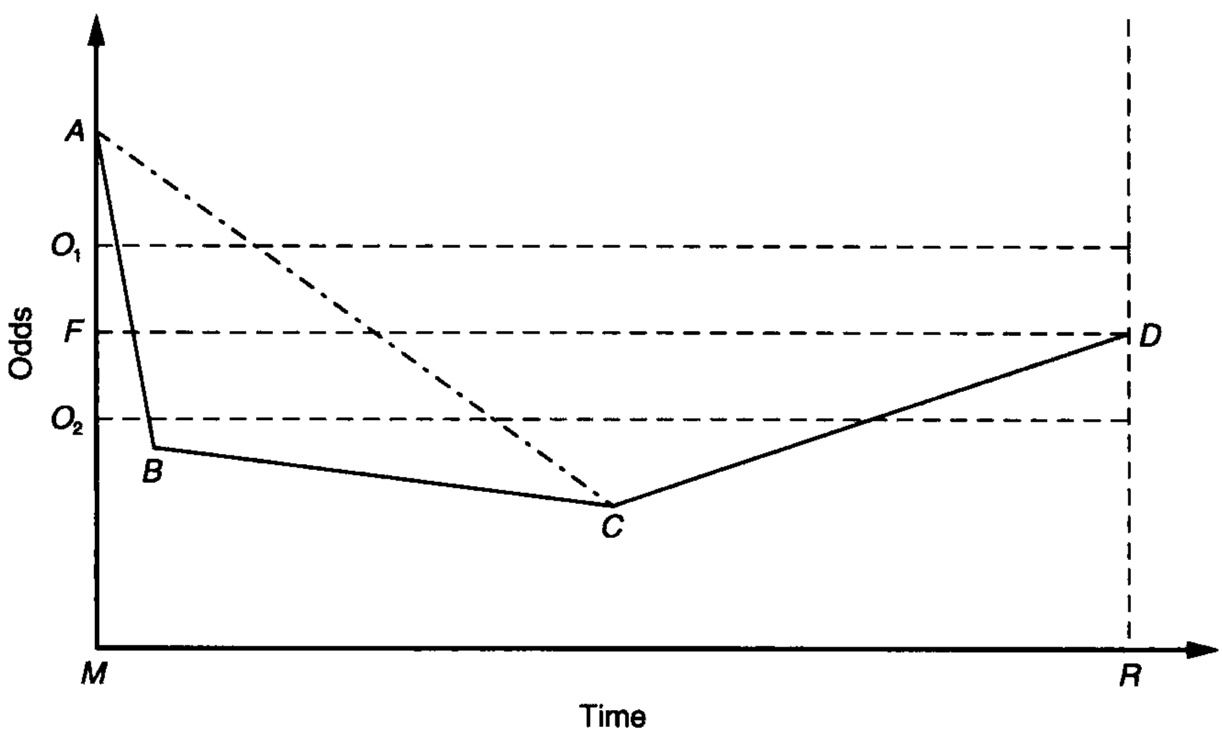


Figure 17.1 Alternative paths of odds of a horse selected by an influential media journalist.

The further condition for semi-strong efficiency is that there should be no arbitrage opportunities between times *M* and *R*. Assume that *F* represents fair odds, and that the public believe that the tipping column has proven profitability. Path *ACD* has an initial segment *AC* during which time there is sustained betting support, net of hedging (laying-off liabilities) on the betting exchanges, to the approximate midpoint between *M* and *R*. Odds decrease at a constant rate over the segment *AC*. Beyond this point in time, segment *CD* indicates a reversal of odds movement, with hedging on the exchanges outweighing further betting-to-win. If the situation in *ACD* is representative, semi-strong inefficiency exists in the sense that there are transparent arbitrage opportunities associated with the segment *AC*.

In contrast, the initial adjustment downwards of odds following publication in *ABCD*, represented by the segment *AB*, is much more rapid after publication, followed by segment *BC* (further bets-to-win outweigh hedging by early bettors), and *CD* (hedging outweighs bets-to-win). *ABCD* still constitutes semi-strong inefficiency in relation to the dynamics of the market, but not to the same degree as *ACD*.

The limiting case is where adjustment to fair odds following publication is instant, after which odds remain constant until *R*, represented by *AF* in Figure 17.1. This limiting case offers no opportunity for profitable arbitrage and fulfils both conditions for semi-strong efficiency in the scenario depicted above.

The precise nature of odds adjustment to tipped horses in reality is in question, but anecdotal evidence suggests that in the case of the most popular newspaper tipping features (typically those published in the *Racing Post*, the leading UK horse race trade newspaper) there is rapid assimilation of this information into odds, which frequently contract substantially following publication, broadly consistent with semi-strong efficiency, but that there is often also a period later in the market when odds extend due to bettors locking in value by laying selections on the exchanges.

The current author is unaware of any empirical studies that have fully addressed the dynamics of markets in relation to intermediate stages in bookmaker or betting exchange markets in relation to horse race betting markets; rather, such studies have been confined to analysis of starting price (*SP*)[3] market data alone; or of comparative static analysis of odds at two or three points in time. For example, morning odds, opening show,[4] and *SP* (Crafts, 1985).

Smith (2003) studied the relationship between the tips of UK racing journalists and the price movements and profitability of racehorses that they select. Greater attention was paid to one tipster than the others, *Pricewise* of the *Racing Post*. The idea behind the column is to identify 'over-lays'; that is, horses whose true chance is understated by the bookmakers' prices – the odds are too big. The column recommends an average of between two and three bets on days featuring high-profile races, sometimes more than one in a race. *Pricewise* has long been portrayed in the media as a feature that leads to significant odds movements, and claims to be a notably profitable newspaper column. In relation to 4,434 horses, Smith measured odds movements from early-morning bookmaker odds to *SP*; the measure used was that employed by Crafts (1985).[5]

Horses in the sample were categorised by tip status, then average returns were calculated based on two staking strategies: unit stake and ticket price to win £1 (the latter calculated as the reciprocal of 1 minus nominal odds probability).

In terms of returns, the performance of such horses was understated by the early-morning odds, indicated by high positive returns at the early odds. The high rates of return to *Pricewise* horses at early-morning odds suggested that the bookmakers' initial appraisal of the chances of these horses was erroneous, and the *Pricewise* assessment superior.

Media tips appeared to have a significant impact on prices from the morning odds to *SP*, suggesting that knowledge of *Pricewise* and nap[6] selections by other journalists are a useful predictor of large contractions in price, especially the former. However, the study also suggested that a significant number of *Pricewise* horses ultimately drifted from early-morning odds to a higher *SP*, and that a small positive return, relative to the overall return to *Pricewise* selections would have been earned on this group at *SP* (approximately 5 per cent). This was very close to the *SP* return for all *Pricewise* horses, suggesting that bettors were able to discriminate between *Pricewise* selections, making judgements on both the validity of the column's arguments supporting the selections, and its claim that they are over-priced. The problem for bettors is *ex ante* identification of these movers prior to price contraction. The positive returns at *SP* concerning such horses were also not statistically significant; therefore, semi-strong efficiency in respect of final odds could not be rejected.

This evidence suggests that media journalists in the sample had a high degree of expertise in regard to such runners relative to the early bookmaker assessment contained in the morning odds, but this could not be translated into statistically significant positive profits unless either assimilation of tips into the market was slow, potentially allowing arbitrage (which anecdotal evidence suggested was not the case); or unless filtering strategies could be employed to isolate those tips which were likely to contract.

In relation to newspaper experts' forecasts of odds (as opposed to forecasts of the winners of races), studies by Snyder (1978a, 1978b) suggested that notional returns to forecasted odds exhibited a greater favourite–longshot bias than final pari-mutuel odds. Losey and Talbott (1980) found that bettor returns could actually be improved by avoiding horses where a comparison of forecast odds and final pari-mutuel dividend implied a positive expected value to wagers struck at the latter odds, suggesting a greater forecasting error for the experts relative to the market. The evidence suggests that it is not possible to devise a profitable trading strategy based on the published odds forecasts of experts.

3 Private handicappers

A further set of studies consider the potential of sophisticated, often computer-intensive, data-processing techniques in formulating profitable wagering strategies based on fundamentals, rather than the more qualitative and rule-of-thumb heuristics typical of newspaper forecasters' decision-making processes. They further differ from the media forecasts insofar as the specific selections arising from these methods are not generally available to the general betting public.

Two landmark articles outlining the methodology for constructing and applying models based on fundamentals to race wagering are considered here. The first, presented by Bolton and Chapman (1986), estimated true probabilities from fundamental data alone. The second, by Benter (1994), builds on the Bolton and Chapman paper by stressing the importance of incorporating the public's estimation of probabilities as evidenced in the odds.

Bolton and Chapman adopted the multinomial logit model (a maximum likelihood technique) of the form:

$$\exp(L) = \prod_{j=1}^{J} p_{jh}, \tag{17.2}$$

where j is the race index, L is the log likelihood function, h is the winner of race j, and

$$p_h = \frac{\exp(v_h)}{\sum_{h=1}^{H} \exp(v_h)} \tag{17.3}$$

where p_h is the estimated true probability of horse h winning race j. Here, v_h is an attribute vector that characterises the fundamental attributes $z=(1, 2,\ldots, n)$, such as distance, going, track configuration, etc. Therefore,

$$v_h = \sum_{n=1}^{N} \Theta_n z_{hn}, \tag{17.4}$$

where θ, a coefficient associated with attribute vectors z, is estimated to show the relative importance of each of the n attributes z for an individual horse h, in explaining the winner in the log likelihood function. The model permits an assessment of the goodness of fit for particular specifications of v_h, through the use of the McFadden (1974) statistic:

$$R^2 = 1 - \frac{L(\Theta - \hat{\Theta})}{L(\Theta = 0)} \tag{17.5}$$

A second statistic, derived from the multinomial logit model, proposed by Wald (1943), was utilised by Bolton and Chapman (1986) and in a number of subsequent empirical studies (a good recent example is Sung *et al.* 2005). The Wald statistic is specified as:

$$-2(L(\Theta = \hat{\Theta}^{(1+2)}) - \left[L(\Theta = \hat{\Theta}^{(1)}) + L(\Theta = \hat{\Theta}^{(2)})\right]), \tag{17.6}$$

where $\theta^{(1)}$ and $\theta^{(2)}$ are the log likelihood values for two separate data sets, and $\theta^{(1+2)}$ is the corresponding value for the two data sets pooled. Using the Wald statistic it is possible to test whether two data subsets have the same parameter specification, as the Wald statistic is distributed as χ^2, with N degrees of freedom, where N is the number of parameters. The Wald statistic can be further exploited to see if one particular maximum likelihood specification explains the information in a data set (in this case the incidence of winners in the sampled races) better than another specification, either based on additional variables z, or using existing z variables specified differently.

The fundamental variables included in the Bolton and Chapman model, in abridged form, were: *Races won as a percentage of lifetime runs*; *average speed rating*; *winnings per race*; *speed rating of last performance*; *percentage and number of winning rides of the jockey*; *weight carried*; *post position*; and *race distance*.

Those familiar with the multifaceted nature of horse racing data will appreciate that the amount of information suggested by the above variables would make this a complex model. Nonetheless, it is equally clear that there are many other omitted explanatory and interaction variables which may be relevant to race outcomes: *pedigree*, *interactions between distance and going*, *track configuration* and so on. Mordin (1993) and Beyer (1994) describe the interplay of factors influencing race outcomes in relation to UK and US horse racing, respectively.

Bolton and Chapman found that the signs of all the estimated coefficients of variables z in their model were in accordance with *a priori* expectations. The most influential predictors in the model, as measured by coefficients θ standardised for standard deviation, proved to be a horse's average speed rating; winnings per race in the current year; and the jockey's percentage of winning rides in the current year.

The authors applied a betting strategy based on estimations of true probabilities derived from their MLE model. This was done by repeated estimation of the model in subsamples comprising three quarters of the 200-race sample, with application of the resulting parameters to derive true probabilities from $v_h = \Sigma \theta z_h$ for each runner in the remaining quarter of races, followed by normalisation to account for in-race competition. The virtue of testing the model out-of-sample in this way is that upward bias of statistical diagnostics through overfitting is avoided.

Bolton and Chapman tested several possible wagering strategies, finding evidence of profitability to a simple unit stake and to a strategy devised by Rosner (1975), which approximates the Kelly criterion (Kelly, 1956; Gottlieb, 1985). Profits, however, only resulted when a minimum estimated true probability filter was applied, restricting bets to horses where the estimated probabilities fell within a range of between approximately 0.17 and 0.22. Estimated low-probability and high-probability winners, conversely, performed poorly, indicating a significant degree of error in the model; this is to be expected as a consequence of the limited number of explanatory variables included.

Benter (1994), who is notable for achieving considerable wagering success with a similar model to that of Bolton and Chapman, suggested that the choice of appropriate fundamental variables is to some extent always incomplete or involves considerable data mining, with the inevitable consequence of forecasting error. Similarly, neural networking, a technique which has been employed in recent years to develop models which efficiently learn from past race results (May, 1998, outlines such a model in some depth), is prone to the statistical bias associated with overfitting of the data.

Benter suggested that the risk of spurious forecasts arising from fundamental handicapping models as a result of data mining can be substantially reduced by including the public's estimates of true probability, introducing market odds as an explanatory variable. He demonstrated that taking this measure decreased the degree of model forecasting error by an appreciable extent, claiming positive returns to a better specified model including odds as a variable.

A further subset of experts whose forecasts of race outcomes are not widely disseminated to the betting public at large are the group of professional and semi-professional bettors whose livelihood is wholly or partially derived from betting their selections, which are generally derived from detailed form analysis. To succeed, their selection and/or money management strategies must be better than the market as a whole, and the evidence suggests that this group is small in number (Nevison, 2008: 109–114); without details of the numbers of bettors who attempt to make a living from betting horses and fail, it is not easy to judge the extent to which those who succeed do so by good luck or sound decision-making and judgement.

From the detailed autobiographies of a number of such professionals (Carroll, 1991; Kuck, 1993; Brohamer, 2000; Beyer, 1993, 1994; Mordin, 1993, 2003; Quinn, 2003; Nevison, 2008, 2009; Veitch, 2009) it is evident that they employ form analysis; horse selection; choice of bet type (win, place, etc.); and money-management heuristics which differ substantially from the statistical methods used by Benter, for example. Much of their analysis revolves around, to coin an American term, 'trip handicapping', which involves the part-qualitative and part-quantitative weighting and measurement of race variables (typically those specified in relation to the Bolton and Chapman model, but in addition with emphasis on race tactics and attention to 'trips', the latter referring to the degree of difficulty or impediment suffered by horses in running in their past races). In addition, trip handicappers often make use of summative form or time ratings as measures of racehorse performance (more on these in a later section).

Evidence from the above autobiographies on the betting methods employed by 'high rollers' (for want of a better shorthand term), suggests that a notable characteristic shared by many of these individuals, defining their relationship to the market, is the tendency to exhibit a confidence in their own judgement bordering on arrogance. Once the horse selection is made, it is common, for example, to find that the high rollers step in to bet again, should the horse's odds drift in the market, on the grounds that if their initial assessment of the horse's chances was correct, the higher odds offer greater value. In this way, the high roller has typically differed from Benter, for example, who incorporated the public assessment contained in the odds into the forecasting model; in addition, the evidence of Crafts (1985) suggests that such a betting strategy is associated with a negative expected value.

From the following quote, it would appear that at least one high roller has modified this view of the significance of market signals, and is struggling to reconcile the value of his own judgement with that of the market:

> I believe that market movements are now more significant than they used to be and that I have made less money than I could have done recently because I have continued to put my tissue[7] before the market.... I should probably give more weight to late changes in market sentiment and keep backing horses when they pass through my tissue price in a positive direction, while resisting the temptation to back those moving through my tissue price in a negative direction. I don't find that easy, because it involves reshaping my concept of value.
>
> (Nevison, 2009: 218–219)

The lack of awareness by a successful professional bettor of market phenomena published by Crafts, for example, as long ago as 1985 casts doubt in the current author's mind on the degree to which historical price data impacts on the decisions made by professionals involved in the markets in question!

4 Bookmakers

Most studies of bookmaker markets indirectly infer the superiority of bookmaker evaluation of probabilities over those of bettors from the existence of persistent negative returns to bettors in aggregate. However, if bettors receive consumption utility from placing wagers in addition to utility from monetary returns, aggregate negative returns do not necessarily imply inferior bettor decisions to those of bookmakers. Furthermore, these studies tell us nothing about the abilities of bettors who choose to refrain from entering the market when they judge that bookmaker prices overstate the true chances of race entrants, or the terms of the transaction are otherwise unfavourable. For example, high deductions from winnings may represent a hurdle rate sufficient to deter skilful bettors. A more comprehensive test of the relative sophistication of bookmakers and bettors in assessing the true chances of a range of outcomes would permit bettors to express alternative prices to bookmakers so that the distribution of revealed preferences of both groups can be observed.

A study by Levitt (2004) evaluated the relative assessments of bookmakers and bettors with reference to data from a handicapping competition based on US National Football League matches. Levitt suggested that a key difference between conventional financial asset markets and betting markets is that in the former the complexity of information affecting the value of assets is such that market makers cannot gain an advantage through superior processing of information to the market as a whole, whereas in the latter, market makers (bookmakers) possess skills in assessing the true chances of various outcomes superior to most bettors, and at least as good as the subset of most skilful bettors.

Levitt suggested that the structural consequences of this differential degree of sophistication are that spot markets equalising supply and demand prevail in conventional financial assets markets, with market makers earning the bid–ask spread, whereas profit-maximising bookmakers set prices to exploit bettor biases, constrained only by the presence of the smaller number of unbiased bettors. Bookmakers therefore earn the equivalent of a bid–ask spread (known as over-round) and an additional return accruing from their exploitation of bettor biases. One consequence of this tendency of bookmakers to act as price makers is that individual books will expose them to positive risk, as bookmakers assume long and short positions exploiting bettor biases.

A disadvantage of the Levitt approach is that, for his data, bookmakers set the terms of the transaction, and bettors respond with a simple decision of whether to bet or not. The most skilful players in this situation may be exercising their talents most effectively in cases where they leave specific games alone, but these decisions are not measured in the Levitt study.

Smith *et al.* (2009) corrected for this omission, in a study which tested the hypothesis that bookmakers display superior skills to bettors in predicting the outcome of sporting events, by using matched data from traditional bookmaking and person-to-person exchanges.[8] The already extensive literature on the economics of auctions (e.g. Klemperer, 1999, 2004) suggests that the decentralised nature of the decision-making processes characteristic of betting exchanges would accomplish the aggregation of dispersed information in a very efficient manner, whereas the bookmaker (however well informed) may fail to match as efficiently the information revealed through such decentralised bidding. Employing a conditional logistic regression model on horse racing data from the United Kingdom, Smith *et al.* found that, in high liquidity betting markets, betting exchange odds have more predictive value than the corresponding bookmaker odds. To control for potential spillovers between the two markets, they repeated the analysis for horses where odds diverged significantly between the two markets. Once again, exchange odds yielded more valuable information concerning race outcomes than the bookmaker equivalents.

The results derived by Smith *et al.* (2009) contrasted with those of Levitt, who found that bookmakers exhibited superior skills in evaluating objective outcomes in the handicapping contest that was the medium for his study. Levitt's methodology meant that the preferences of the most skilled or informed bettors might not be revealed if they decided that the terms of the wagers set by bookmakers were unfavourable and, in consequence, chose not to trade. In contrast, the betting exchanges offer opportunities for these bettors to trade, which are not available in bookmaker markets. For example, skilled traders, insiders and bettors seeking hedging opportunities are all able to lay odds on the exchanges which may, as a result, more accurately reflect the chances of the horses concerned than those offered by bookmakers. In these circumstances it is plausible that the proportion of turnover attributable to casual bettors will be lower in the exchanges than in bookmaker markets, with a consequent tendency for odds to reflect more closely the corresponding objective probabilities. Differences in the nature of traders and trading activities may therefore explain the greater relative efficiency of the exchanges in reflecting objective outcome probabilities observed.

The Smith *et al.* findings are consistent with the empirical evidence suggesting that decentralised markets are efficient predictors. They also resonate with anecdotal evidence that bookmakers increasingly employ betting exchange data as the basis for their odds (Nevison, 2009: 59).

The empirical evidence suggests that *Betfair*'s claim that their odds on average exceed those of bookmakers by approximately 20 per cent is credible (based on matched bookmaker and exchange data employed by Smith *et al.*, 2005). When odds structures are examined more closely, however, bookmakers starting prices are relatively efficient predictors of high-probability winners, and competitive with exchange odds for this class of runner (Smith and Vaughan Williams, 2008; Nevison, 2008: 165–166) due to the negation of a more marginal exchange odds advantage by the commission charged on winnings by the exchanges (typically 5 per cent on *Betfair*, while bookmakers make no explicit deduction from winnings). In contrast, bookmaker odds systematically overstate the chances of longshots by a large margin, whereas exchange odds tend on average to be close to their true probability of winning: hence the favourite–longshot bias observed in bookmaker markets.

Shin (1991, 1992, 1993) offered a behavioural explanation of this bias, whereby bookmakers seek to preserve margins as a consequence of insiders who are in a position to exploit monopoly knowledge of the true winning chance of a specific horse in a race, by depressing the odds against longshots from their true values. Shin derived an equilibrium profit-maximising price structure solving this adverse selection problem, offering a plausible explanation for the favourite–longshot bias based on the presence of insiders. The model enables empirical estimation of the proportion of betting turnover attributable to insider trading in individual races or groups of races; furthermore, Jullien and Salanié (1994) and Cain *et al.* (2001) show that the Shin model can be employed to derive efficient estimates of objective probabilities for individual runners in a race, by adjustment of the odds for bias, from a knowledge of the nominal odds and the number of runners in the race. This suggests that bookmakers employ sophisticated strategies to protect their interests in the face of insider activity.

In practice, Shin-adjusted probabilities always imply that bookmaker nominal odds overstate their true values (by a small margin for high-probability runners and a large margin for low-probability runners). Trading strategies based on these efficient, adjusted-for-bias, Shin probabilities, therefore, can only be achieved if an alternative odds set exists to permit arbitrage – Vaughan Williams and Paton (1997) find such quasi arbitrage opportunities – or by permitting traders to lay horses on the exchanges at odds less than their estimated Shin values.

Despite the popularity of the Shin model in recent years as a means of deriving a proxy measure of the favourite–longshot bias (by the estimation of 'Shin's z', a coefficient of the model), its behavioural underpinning, and consequently its validity as an explanation of bias, has been called into question. Schnytzer and Shilony (2003) show that the favourite–longshot bias can exist if Shin's z is equal to zero (implying the absence of insiders in the market); Peirson and Smith

(2010) also derive a theoretical model of bookmaker odds setting which shows that there is little difference between the impact of the operation of insiders and expert gamblers on the prices set by bookmakers.

5 New evidence on market assimilation of media forecast information

In his sampled UK horse races, Crafts (1985) suggested a strong association between high notional returns at opening odds accruing to wagers on horses significantly contracting in odds to *SP* and insider betting, but did not seek confirmation of the nature of this association by establishing the amount of public information available on the constituent runners.

The present author conducted a similar study of price movements in relation to horses top-rated by the *Racing Post* form ratings, for a large sample of handicap and high-class stakes horse races run in the United Kingdom from August 2009 to July 2010, to establish whether similar results would be achieved in relation to horses with ostensibly strong, publicly available, form credentials. The *Racing Post* ratings were chosen because they are widely published on the morning of the races, and are formulated according to established and widely accepted handicapping principles whereby ratings are determined by the relative finishing positions of horses relative to each other in a race, with adjustments for weight carried, the number of lengths horses finish behind the winner and the race distance. The ratings have the additional merit of being widely recognised as authoritative measures of racehorse ability (Nevison, 2008, 2009).

Opening show odds were acquired from *sportinglife.com* for all *Racing Post* top-rated horses in the race sample, where the opening odds were 8/1 or less. The 8/1 cut-off point was chosen (1) to minimise the likelihood of returns reflecting the favourite–longshot bias,[9] and (2) to filter out spurious selections arising from the inability of the ratings to allow for current race conditions (track, racing surface, race distance and so on) which may detract from a horse's performance.

The author employed a measure of odds movement from opening show to *SP* weighted by odds, suggested by Law and Peel (2002), in preference to the Crafts measure, which does not allow for the greater trading volumes required to move odds at lower odds. Law and Peel employed the measure, p_m, such that:

$$p_m = \log\left(\frac{1}{1-p_1}\right) - \log\left(\frac{1}{1-p_2}\right), \tag{17.7}$$

where, for an individual runner in a race, p_1 and p_2 are the odds probabilities derived from, for example, starting odds and forecast odds, respectively. Unlike the Crafts ratio, p_m weights price movements from initially low odds with greater emphasis than those from initially high odds, reflecting the greater trading volumes required to cause odds to change at low odds. For illustrative purposes, Table 17.1 indicates the divergence for different levels of odds required to yield specific values of p_m.

Table 17.1 Odds contractions to *SP* corresponding to selected values of p_m

Initial odds	*SP odds, yielding specified pm values*				
	$p_m=0.01$	$p_m=0.02$	$p_m=0.04$	$p_m=0.08$	$p_m=0.15$
3 to 1	2.88	2.78	2.58	2.25	1.82
4 to 1	3.81	3.63	3.32	2.82	2.21
5 to 1	4.72	4.46	4.02	3.33	2.54
6 to 1	5.61	5.26	4.67	3.79	2.81
7 to 1	6.48	6.03	5.28	4.20	3.05
8 to 1	7.34	6.77	5.85	4.57	3.26

Notes

1 *pm* is measured as in Equation (17.7) – see also accompanying narrative.

2 All odds expressed to a one-unit stake, e.g. '2.88 to1', '6.77 to 1' and so on.

Average returns (total profit/loss divided by total stake), by banded values of p_m, were computed for top-rated horses whose odds contracted during the betting period, at a level unit stake. Similarly, returns were calculated for the subset of top-rated horses drifting in the market. Edelman (2001) suggests that horses whose odds are the same at the beginning and end of the betting period may constitute 'hidden overlays'. On these grounds, horses whose p_m value was zero were included in the contractor group. The above procedure was then repeated for odds contractors and drifters in the non-top-rated category.

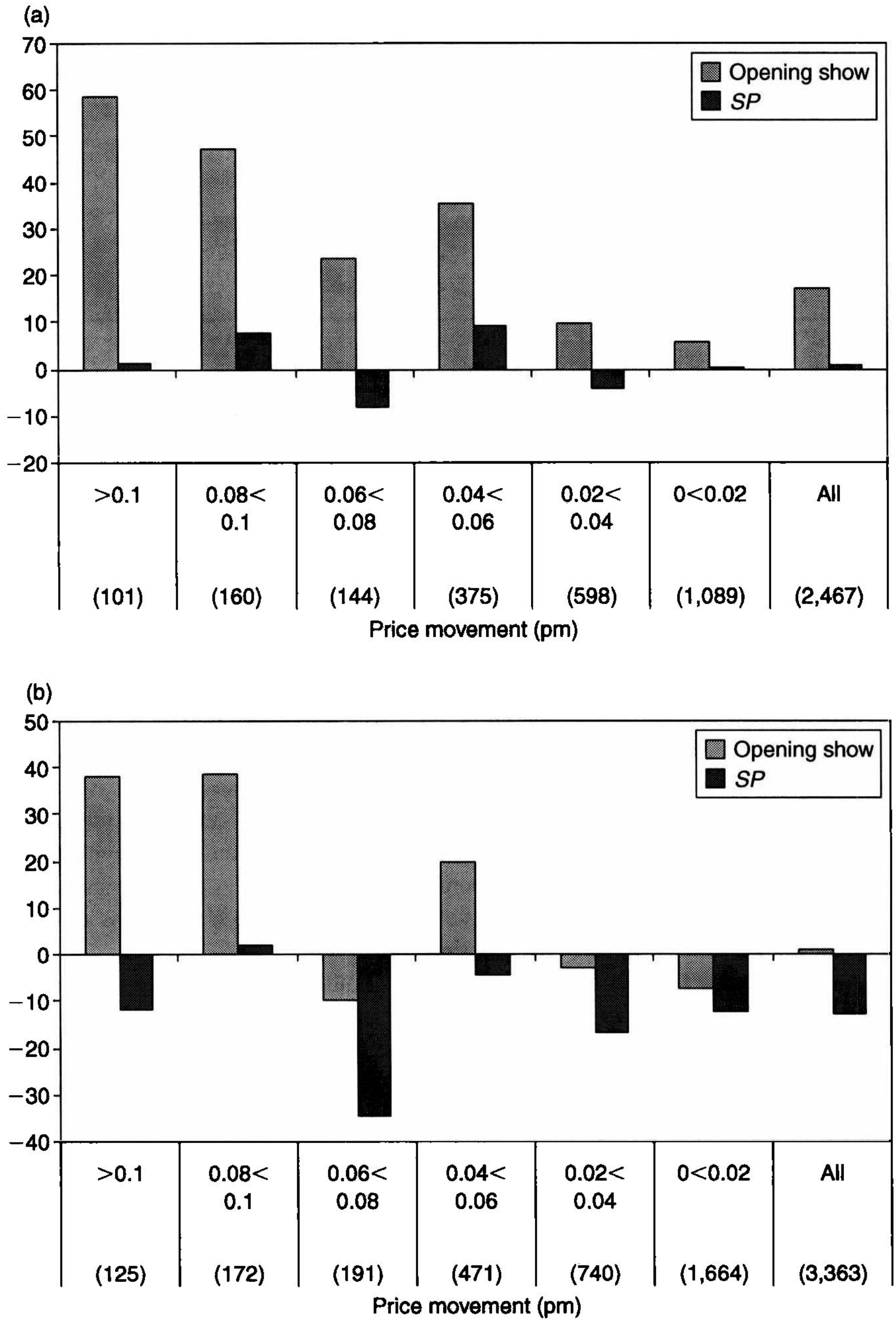
The average returns at opening show odds to a unit stake placed on all tipped horses was 2.65 per cent, close enough to zero to suggest that the hypothesis that bookmakers accurately assess the chances of top-rated horses as a class cannot be rejected. The corresponding return at *SP* was 0.44 per cent, suggesting the market is semi-strongly efficient at final odds. In contrast, returns to level stakes wagered on all non-top-rated horses were strongly negative at opening show and *SP*, suggesting that the bookmakers overestimate their true chances, as do subsequent market forces in the betting period.

As with Crafts (1985), we can learn more about the efficiency of the market from its dynamics from opening show to *SP*.

Figures 17.2(a–d) reveal the returns, categorised by p_m values, to a level stake wagered on top-rated contractors, top-rated drifters, non-top-rated contractors and non-top-rated drifters.

Figures 17.2(a) and 17.2(c) suggest that *SP* efficiently assimilates information about the chances of runners across the p_m categories; the returns for >0.1 and 0.08<0.1 drifters are atypical, with small numbers of observations. It is notable from Figure 17.2(a) that notional returns to top-rated contractors at opening show indicate that bookmakers underestimate the chances of this subset of form horses, as opposed to the market as a whole, which appears able to isolate these horses and accurately drive their odds down to *SP*.

In contrast, bookmakers accurately assess the chances of those non-top-rated horses at opening show which subsequently contract, with the exception of the biggest odds movers, whereas the market at *SP* over-bets them by some margin,



(a)
70
60
50
40
30
20
10
0
−10
−20
Opening show
SP
>0.1
0.08< 0.1
0.06< 0.08
0.04< 0.06
0.02< 0.04
0<0.02
All
(101)
(160)
(144)
(375)
(598)
(1,089)
(2,467)
Price movement (pm)
(b)
50
40
30
20
10
0
−10
−20
−30
−40
Opening show
SP
>0.1
0.08< 0.1
0.06< 0.08
0.04< 0.06
0.02< 0.04
0<0.02
All
(125)
(172)
(191)
(471)
(740)
(1,664)
(3,363)
Price movement (pm)

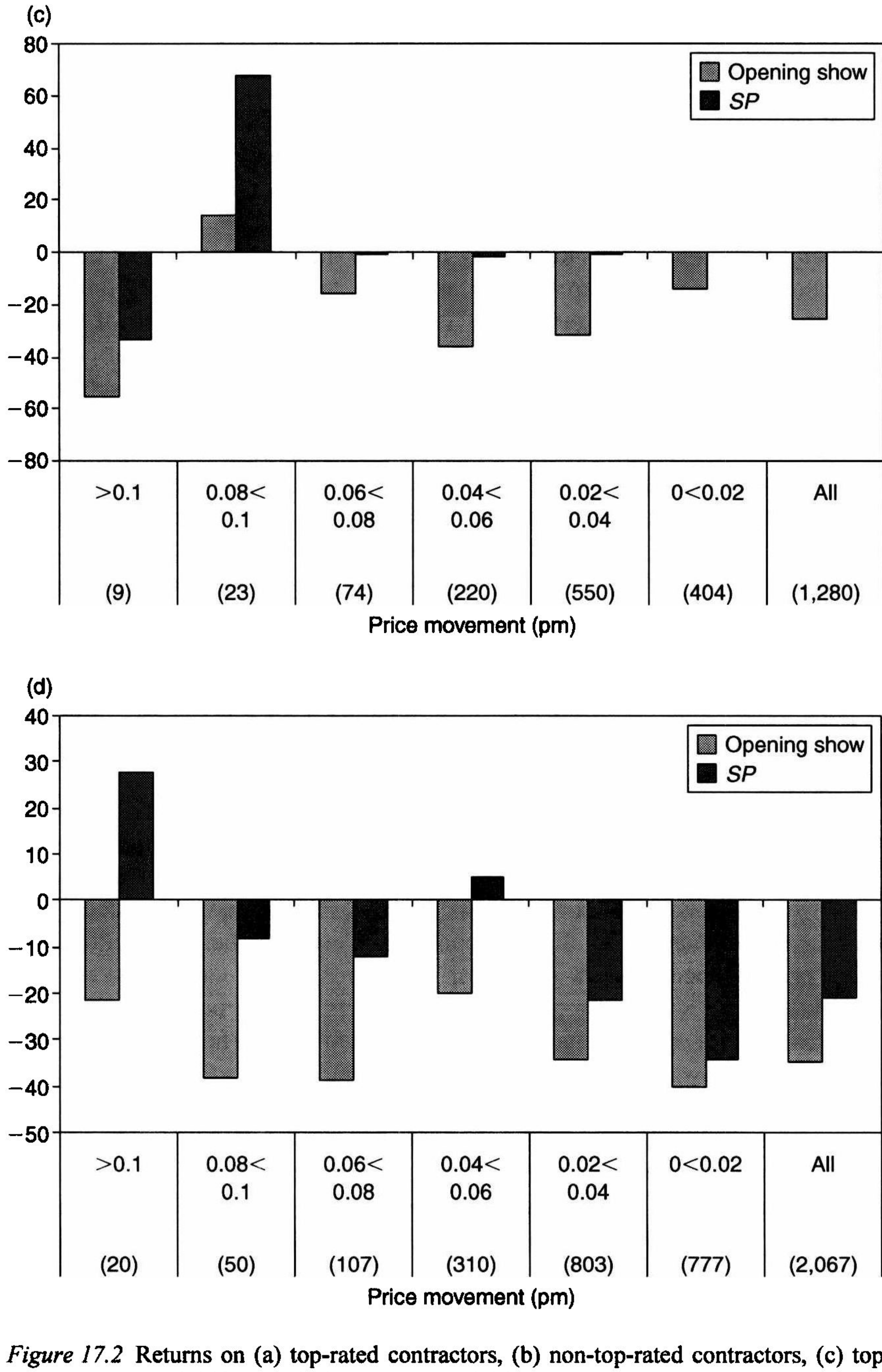


Figure 17.2 Returns on (a) top-rated contractors, (b) non-top-rated contractors, (c) top-rated drifters and (d) non-top-rated drifters.

evidenced by strongly negative returns. Returns to non-top-rated drifters, indicated in Figure 17.2(d), suggest that these horses are over-rated at both opening show and *SP*; while the market as a whole successfully judges that these horses are over-estimated by bookmakers, odds do not subsequently drift sufficiently to reflect their true chance.

The evidence on non-top-rated horses – that contractors and drifters 'overshoot' and 'under-shoot' their objective odds values respectively – contrasts with the top-rated horses, which the market clears efficiently to *SP*. This general overestimation of non-top-rated horses may be due to the preponderance of horses in this category for which expectations are high, and public form is limited in amount and quality but suggestive of improvement to the standard of today's race. Some of these horses may, for example, be easy recent winners of previous lower-class races, judged to have the scope to make significant improvement. In addition, relatively unexposed horses from powerful stables; from stables with a reputation for gambling; and horses ridden by top jockeys but with otherwise flimsy form credentials, frequently fall into this price range and become the medium of strong betting by the public, and/or conservative odds setting by bookmakers, justified more by potential than previous performance.

The marked increase in positive notional returns at opening show odds accruing to wagers on contractors as p_m increases, for both top-rated and non-top-rated horses, indicates that knowledge of the magnitude of odds movements holds valuable information in addition to that inferred from a knowledge of the direction of price movements. The difficulty in formulating a trading strategy to exploit these notional profits at opening show is to know at the commencement of the betting period which horses will contract to *SP* and by how much.

Overall, it would appear from the evidence of this study that media selections, drawn from those horses with relatively robust public form, are in aggregate correctly valued by bookmakers. The chances of the subset of horses in this category which subsequently contract in odds, however, are underestimated by the opening show odds, either through bookmaker error, or because bettors who wager prior to opening show (when early-morning odds are available) are poor at evaluating the winning probabilities of runners. This group of horses (top-rated contractors) are subsequently correctly valued at *SP*. These results are consistent with Figlewski's finding that the wagers placed in the final fraction of the betting period are the most informed.

The returns distributions of top-rated horses appear to be consistent with Crafts (1985), yet Crafts associated strong contractors with insider activity, which is not the case here as much is known about the intrinsic ability of the horses in question; we have already seen that this association is disputed in recent studies.

The new evidence presented in this chapter, and the earlier results in relation to the *Pricewise* column, show that odds contractions attributed to insiders by Crafts, yielding high notional profits at the initial odds values, are more closely associated with influential tipsters and ratings compilers' selections of well-exposed horses with established publicly known form credentials than with the unknown quantities which one might expect to be the focus of insider trading. The clear difference in returns between subsets of horses whose odds contract and drift respectively in these studies, however, does allow for insider activity associated with privileged knowledge of the current condition of otherwise publicly exposed horses, or marginal influences on their performance not revealed by past form.

6 Conclusions

This chapter has considered evidence on the degree to which key groups of experts in relation to UK horse race betting markets (media forecasters, private handicappers and bookmakers) exercise informative judgement in relation to race fundamentals and the extent to which the market efficiently assimilates their analysis.

Bookmaker starting prices hold more predictive power than media morning odds forecasts. In the case of influential newspaper tipsters, however, the picture is more complex. The evidence suggests that starting prices appear to exhibit semi-strong efficiency in relation to the selections of influential newspaper tipsters, with a negative expected value or positive but insignificant returns.

Bookmakers morning odds and on-course opening show odds exhibit semi-strong efficiency insofar as in aggregate they do not permit profitable trading strategies at their initial values. However, the market successfully identifies initial pricing errors in relation to individual runners in the betting period from the morning/opening show to race time, evidenced by the existence of high positive notional returns to market contractors at morning/opening show bookmaker odds, and strongly negative notional returns at opening odds on horses which subsequently drift in the market.

Recent studies speculating on the origins of favourite–longshot bias suggest that bookmakers are either skilful at exploiting bettor biases (Levitt, 2004) or solving the adverse selection problem arising from the presence of insiders (Shin, 1991, 1992, 1993); however, the evidence on which explanation of the bias holds most empirical validity is equivocal.

It was suggested above that the evidence on media forecast casts doubt on Crafts' unqualified inference of insider activity associated with marked odds movers, as the former exhibit similar returns properties. At the very least, a more nuanced understanding of different types of insider knowledge is required; as Peirson and Smith (2010: 990) observe, there is a danger that policy and regulation punish virtuous market assimilation of public information, rather than penalise undesirable insider activities.

The evidence suggests that experts – influential media 'tipsters', private handicappers and bookmakers – exercise decision-making skills which ostensibly contribute to market efficiency, depending on context. The information contained in their evaluations would equally appear to be significantly enhanced by factoring in market fluctuations and final odds. In addition, with or without market signals, all the evidence suggests that it is difficult for bettors to execute profitable trading strategies based on the publicly available information derived from the decisions of the experts reviewed.

As Benter acknowledges, the tenets of the efficient-markets hypothesis suggest that, even if experts can establish an edge over the market, the greater challenge is to maintain that advantage when the price of success is the negation of profits by market assimilation of the information and analytical processes underlying their wagers (Benter, 1994). The betting exchanges and associated internet-based technologies are changing the betting landscape in ways which render today's expertise redundant tomorrow. A similar review written in the near future will no doubt consider new subsets of experts: traders rather than bettors; exchanges rather than market makers; and other agents as yet unknown. Whatever the nature of tomorrow's experts, the market will remain central to the fortunes of those who act on the judgement of the experts.

Notes

1 'Morning line' is a term used for the set of forecast odds compiled by analysts, most commonly employed by trade or other newspapers, who study past form and related information to estimate the chances of each horse in a race.
2 Betting exchanges are online many-to-many double auctions for bettors, offering a trading platform from which clients can bet-to-win or lay-to-lose with other clients; they are the e-bay of the betting world. *Betfair* is the leading betting exchange, with an approximate 90 per cent share of the exchange market – see Smith and Vaughan Williams (2008) for further information on the functionality of betting exchanges.
3 The starting price, or *SP*, is a unique odds value for each horse determined by official on-course odds inspectors at which winning bets are settled in the absence of a specified fixed-odds value agreed between bookmaker and bettor.
4 In the UK, opening show refers to the initial on-course odds advertised by bookmakers in the betting period immediately preceding the race; typically, the on-course market opens ten minutes before the advertised race time. Odds fluctuations from opening show to the commencement of the race are distributed by the Press Association, to be published online at sites such as *sportinglife.com*, and to off-course betting shops, to facilitate fixed-odds betting off-course. Where fixed odds are not agreed, *SP* applies.
5 Crafts employed the ratio of early-morning newspaper forecast odds to opening show (and separately to starting price) to measure movement for horses contracting (odds reducing); and the inverse of this ratio for horses extending (odds increasing). The results of the Smith study were similar when repeated using an alternative measure of odds movement, also employed later in this chapter.
6 The 'nap' is the horse considered by the journalist to be the best bet of the day.
7 The 'tissue' is the trip handicapper's term for his *a priori* odds estimation for the runners in a race based on his own evaluation of each horse's objective probability of winning.
8 While acknowledging that betting exchange prices may be in part attributable to bookmakers managing their liabilities, the authors argued that significant deviations in values between exchange and bookmakers arise from the actions of non-bookmaker traders.
9 The consensus of studies measuring the favourite–longshot bias show that when bookmakers' average margin per race (known as over-round), or deductions from winnings in pari-mutuel markets, are added back in to returns, the cross-over point

between average and marginal bettor returns is equal between about 5/1 and 8/1, depending on the market sampled. The author has independently confirmed a value of 8/1 in relation to 12 years of UK bookmaker odds dating through 1997–2008.

References

Ball, R. and Brown, P. (1968) 'An Empirical Evaluation of Accounting Income Numbers'. *Journal of Accounting Research*, 6, 159–178.

Benter, W. (1994) 'Computer Based Horse Race Handicapping and Wagering Systems: A Report', in Hausch, D.B., Lo, V.S.Y. and Ziemba, W.T. (eds), *Efficiency of Racetrack Betting Markets*, London: Academic Press, pp. 511–526.

Beyer, A. (1993) *Beyer On Speed*, New York: Houghton Mifflin.

Beyer, A. (1994) *The Winning Horseplayer*, Boston: Houghton-Mifflin.

Bird, R. and McCrae, M. (1987) 'Tests of the Efficiency of Racetrack Betting Using Bookmaker Odds', *Management Science*, 33 (12), 1552–1562.

Bolton, R.N. and Chapman, R.G. (1986) 'Searching for Positive Returns at the Track: A Multinomial Logit Model for Handicapping Horse Races', *Management Science*, 32, 1040–1059.

Brohamer, T. (2000) *Modern Pace Handicapping*, New York: Daily Racing Form Press.

Cain, M., Law, D. and Peel D.A. (2001) 'The Incidence of Insider Trading in Betting Markets and the Gabriel and Marsden Anomaly', *The Manchester School*, 69 (2), 197–207.

Carroll, C. (1991) *Handicapping Speed: The Thoroughbred and Quarter Horse Sprinters*, New York: The Lyons Press.

Crafts, N.F.R. (1985) 'Some Evidence of Insider Knowledge in Horse Race Betting in Britain', *Economica*, 52, 295–304.

Edelman, D. (2001) *The Compleat Horseplayer*, Bulli, Australia: De Mare Consultants.

Fama, E.F. (1970) 'Efficient Capital Markets: A Review of Theory and Empirical Work', *Journal of Finance*, 25 (2), 383–417.

Figlewski, S. (1979) 'Subjective Information and Market Efficiency in a Betting Market', *Journal of Political Economy*, 87 (1), 75–88.

Gottlieb, G. (1985) 'An Optimal Betting Strategy for Repeated Games', *Journal of Applied Probability*, 22, 787–795.

Jullien, B. and Salanié, B. (1994) 'Measuring the Incidence of Insider Trading: A Comment on Shin', *Economic Journal*, 104, 1418–1419.

Kelly, J.L., Jr. (1956) 'A New Interpretation of Information Rate', *Bell System Technical Journal*, 35 (7), 917–926.

Klemperer, P. (1999) 'Auction Theory: A Guide to the Literature', *Journal of Economic Surveys*, 13(3), 227–286.

Klemperer, P. (2004) 'Auctions: Theory and Practice', *Economics Papers 2004-W09, Economics Group*, Nuffield College, University of Oxford.

Krauss, A. and Stoll, H. (1972) 'Price Impacts of Block Trading on the New York Stock Exchange', *Journal of Finance*, 27, (2), 210–219.

Kuck, H. (1993) *Winner's File: The Logic Behind Successful Handicapping*, New York: William Morrow.

Law, D. and Peel, D.A. (2002) 'Insider Trading, Herding Behaviour, and Market Plungers in the British Horse Race Betting Market', *Economica*, 69, 327–338.

Levitt, S. (2004) 'Why are Gambling Markets Organised so Differently from Financial Markets?', *Economic Journal*, 114 (4), 223–246.

Losey, R.L. and Talbott, J.C., Jr. (1980) 'Back on the Track with the Efficient Markets Hypothesis', *Journal of Finance*, 35, 1039–1043.

McFadden, D. (1974) 'Conditional Logit Analysis of Qualitative Choice Behaviour', in Zarembka, P. (ed.), *Frontiers in Econometrics*, New York: Academic Press.

May, P. (1998) *Forecasting Methods For Horseracing*, Newbury: Raceform Ltd.

Mordin, N. (1993) *Betting for a Living*, London: Rowton Press.

Mordin, N. (2003) *Mordin On Time*, Nantwich: Aesculus Press.

Nevison, D. with Ashworth, D. (2008) *A Bloody Good Winner*, Newbury: Highdown.

Nevison, D. with Ashworth, D. (2009) *No Easy Money: A Gambler's Diary*, Newbury: Racing Post.

Patell, J.M. and M.A. Wolfson (1984) 'The Intraday Speed of Adjustment of Stock Prices to Earnings and Dividend Announcements', *Journal of Financial Economics*, 27, 223–252.

Peirson, J. and Smith, M.A. (2010) 'Expert Analysis and Insider Information in Horse Race Betting: Regulating Informed Market Behavior', *Southern Economic Journal*, 76 (4), 976–992.

Quinn, J. (2003) *The Best of Thoroughbred Handicapping: Leading Ideas and Methods*, New York: Daily Racing Form Press.

Rosner, B. (1975) 'Optimal Allocation of Resources in a Pari-mutuel Setting', *Management Science*, 21 (9), 997–1006.

Schnytzer, A. and Shilony, Y. (2003) 'Is the Presence of Inside Traders Necessary to Give Rise to a Favorite–Longshot Bias?', in Vaughan Williams, L. (ed.) *The Economics of Gambling*, London: Routledge.

Shin, H.S. (1991) 'Optimal Betting Odds Against Insider Traders', *Economic Journal*, 101, 1179–1185.

Shin, H.S. (1992) 'Prices of State Contingent Claims with Insider Traders, and the Favourite–Longshot Bias', *Economic Journal*, 102, 426–435.

Shin, H.S. (1993) 'Measuring the Incidence of Insider Trading in a Market for State-contingent Claims', *Economic Journal*, 103, 1141–1153.

Smith, M.A. (2003) 'The Impact of Tipster Information on Bookmakers' Prices in UK Horse-race Markets', in Vaughan Williams, L. (ed.) *The Economics of Gambling*, London: Routledge.

Smith, M.A. and Vaughan Williams, L. (2008) 'Betting Exchanges: A Technological Revolution in Sports Betting', in Ziemba, W.T. (ed.) *Handbook of Sports and Lottery Markets*, Toronto: Elsevier.

Smith, M.A., Paton, D. and Vaughan Williams, L. (2005) 'An Assessment of Quasi-Arbitrage Opportunities in Two Fixed-Odds Horse-Race Betting Markets', in Vaughan Williams, L. (ed.) *Information Efficiency in Financial and Betting Markets*, Cambridge: Cambridge University Press.

Smith, M.A., Paton, D. and Vaughan Williams, L. (2009) 'Do Bookmakers Possess Superior Skills to Bettors in Predicting Outcomes?', *Journal of Economic Behaviour and Organisation*, 71 (2), 539–549.

Snyder, W.W. (1978a) 'Horse Racing: Testing the Efficient markets Model', *Journal of Finance*, 33 (4), 1109–1118.

Snyder, W.W. (1978b) 'Decision-making with Risk and Uncertainty: The Case of Horse Racing', *American Journal of Psychology*, 91 (2), 201–209.

Sung, M., Johnson, J.E.V. and Bruce, A.C. (2005) 'Searching for Semi-strong Inefficiency in the UK Racetrack Betting Market', in Vaughan Williams, L. (ed.) *Information Efficiency in Financial and Betting Markets*, Cambridge: Cambridge University Press.

Vaughan Williams, L. (2000) 'Can Forecasters Forecast Successfully? Evidence from UK Betting Markets', *Journal of Forecasting*, 19, 505–513.

Vaughan Williams, L. and Paton, D. (1997) 'Why is there a Favourite–Longshot Bias in British Racetrack Betting Markets?', *Economic Journal*, 107, 150–158.

Veitch, P. (2010) *Enemy Number One: The Secrets of the UK's Most Feared Professional Punter*, Newbury: Racing Post Books.

Wald, A. (1943) 'Tests of Statistical Hypotheses Concerning Several Parameters, when the Number of Observations is Large', *Transactions of the American Mathematical Society*, 54, 426–482.

Index

Page numbers in *italics* denote tables, those in **bold** denote figures.

{
"filename": "UFJFRElDVElPTiBNQVJLRVRTX1RIRU9SWSBBTkQgQVBQTElDQVRJT05TXzQwNDI0MDEwLnppcA==",
"filename_decoded": "PREDICTION MARKETS_THEORY AND APPLICATIONS_40424010.zip",
"filesize": 35328018,
"md5": "c297d2b33edcba3e9847420384b32eef",
"header_md5": "dab5cd6451ef63e1fef0a470f8e1d5f0",
"sha1": "7fad3bd664940c0d5cbaacbf3cef43d8bb4ecd11",
"sha256": "3578f99c2d753419be961a7e4d879828cb07e97ade15f0d62a5a29c781ab4cd5",
"crc32": 532640650,
"zip_password": "",
"uncompressed_size": 36427406,
"pdg_dir_name": "PREDICTION MARKETS_THEORY AND APPLICATIONS_40424010",
"pdg_main_pages_found": 264,
"pdg_main_pages_max": 264,
"total_pages": 282,
"total_pixels": 1315677658,
"pdf_generation_missing_pages": false
}