Experimental testing of old and new hypotheses in economics

Experimental testing of old and new hypotheses in economics

Eskil Forsell

Dissertation for the Degree of Doctor of Philosophy, Ph.D., in Economics
Stockholm School of Economics, 2016

Experimental testing of old and new hypotheses in economics

ISBN 978-91-7731-014-3 (printed)
ISBN 978-91-7731-015-0 (pdf)

This book was typeset by the author using LaTeX.

Front cover illustration:

Back cover photo:
Nicklas Gustafsson, Arctistic, 2014

Printed by:
Ineko, Göteborg, 2016

Keywords:
Prediction markets, decision markets, replications, information aggregation, game theory, laboratory experiment, observational learning, regression discontinuity, instrumental variable analysis.

To B, who kept me on my toes.

Foreword

This volume is the result of a research project carried out at the Department of Economics at the Stockholm School of Economics (SSE).

This volume is submitted as a doctor's thesis at SSE. In keeping with the policies of SSE, the author has been entirely free to conduct and present his research in the manner of his choosing as an expression of his own ideas.

SSE is grateful for the financial support provided by the Jan Wallander and Tom Hedelius Foundation which has made it possible to fulfill the project.

Göran Lindqvist
Director of Research
Stockholm School of Economics

Richard Friberg
Professor and Head of the
Department of Economics
Stockholm School of Economics

Acknowledgements

Firstly, I would like to thank Dr. Fredrik Sävje for unknowingly inspiring me to apply for the PhD program in economics. Without him there would be no thesis.

I also thank my supervisors, Prof. Jörgen Weibull and Prof. Magnus Johannesson for providing helpful feedback when needed and always asking the right questions to inspire me to always go one step further.

Additionally, I am grateful to Prof. Colin Camerer, Assoc. Prof. Anna Dreber Almenberg and Assoc. Prof. Per Engström for suggesting interesting research topics and allowing me to join them in efforts to answer some of the questions that those topics gave rise to.

Finally, I also thank my fellow PhD students Adam Altmejd, Emma Heikensten and Siri Isaksson, who were great sources for comments, laughs and bewilderment (in almost equal proportions).

Stockholm, August 24, 2016

Eskil Forsell

Chapter summaries

Evaluating Replicability of Laboratory Experiments in Economics

The reproducibility of scientific findings has been called into question. To contribute data about reproducibility in economics, we replicate 18 studies published in the *American Economic Review* and the *Quarterly Journal of Economics* in 2011-2014. All replications follow predefined analysis plans publicly posted prior to the replications, and have a statistical power of at least 90% to detect the original effect size at the 5% significance level. We find a significant effect in the same direction as the original study for 11 replications (61%); on average the replicated effect size is 66% of the original. The reproducibility rate varies between 67% and 78% for four additional reproducibility indicators, including a prediction market measure of peer beliefs.

Trading performance in prediction markets with different structures

This paper presents preliminary evidence on how researchers in the field of psychology judge the replicability of the 28 effects replicated in the Many Labs 2 project. We use individual surveys in combination with prediction markets to elicit beliefs about two replication success metrics — whether the estimated effect in the replication study is statistically significant, and what the ratio between the original and replicated effect size is. We find that survey answers and final market prices are very highly correlated for the binary measure suggesting that the prediction markets provide little additional value, but that the correlation is lower for the effect size measure.

The impact of decision rules on the predictive accuracy of decision markets

An appealing prospect of prediction markets is that their estimates of how likely future events are to occur can be used as inputs when making a decision. As prediction markets used in this way help guide decisions, they are called decision markets. These have stricter requirements on how scoring rules (payment schemes) must be specified to guarantee that traders are incentivized to trade according to their beliefs. They also require that decision rules (the link between market outcomes and what decision is taken) is specified in certain ways. We let participants trade on hypothetical markets using three different combinations of the rules to explore how the predictive accuracy of the markets is affected. Our main finding is that the decision markets perform worse than traditional prediction markets — likely due to their increased complexity — but that there is little impact of the specific rules used.

Gamelab: An online game-theory laboratory

The *Gamelab* platform offers a novel and easy way to perform experiments in game theory. Its options are flexible enough to allow for a wide range of experiments. It is particularly well designed for play against anonymous and randomly drawn opponents. Thanks to its responsive design it can be used on almost any device with internet access. We here report the implementation of experiments in two different settings. In both settings, the subjects were given data about past aggregate play of the same game, thus giving them the possibility for social learning how to play. This platform thus provides a tool to test non-cooperative solution concepts.

Demand effects of consumers' stated and revealed preferences

Knowledge of how consumers react to different signals is fundamental to understanding how markets work. The modern electronic marketplace has revolutionized the possibilities for consumers to gather detailed information about products and services before purchase. Specifically, a consumer can easily – through a host of online forums and evaluation sites – estimate a product's quality based on either i) what other users say about the product (*stated preferences*) or ii) how many other users that have bought the product (*revealed pref-*

erences). In this paper we compare the causal effects on demand from these two signals based on data from the biggest marketplace for Android apps, Google play. This data consists of daily information, for 42 consecutive days, of more than 500 000 apps from the US version of Google play. Our main result is that consumers are much more responsive to other consumers' revealed preferences, compared to others' stated preferences. A 10 percentile increase in displayed average rating only increases downloads by about 3 percent, while a 10 percentile increase in displayed number of downloads increases downloads by about 25 percent.

Contents

Appendixes

Bibliography

Abeler, J., Falk, A., Goette, L., & Huffman, D. (2011). Reference points and effort provision. *The American Economic Review*, *101*(2), 470–492.

Abramowicz, M. (2004). Information markets, administrative decisionmaking, and predictive cost-benefit analysis. *The University of Chicago Law Review*, (pp. 933–1020).

Almenberg, J., Kittlitz, K., & Pfeiffer, T. (2009). An experiment on prediction markets in science. *PLoS ONE*, *4*(12), e8500.

Alter, A. L., Oppenheimer, D. M., Epley, N., & Eyre, R. N. (2007). Overcoming intuition: metacognitive difficulty activates analytic reasoning. *Journal of Experimental Psychology: General*, *136*(4), 569.

Ambrus, A., & Greiner, B. (2012). Imperfect public monitoring with costly punishment: An experimental study. *The American Economic Review*, *102*(7), 3317–3332.

Anderson, C., Kraus, M. W., Galinsky, A. D., & Keltner, D. (2012). The local-ladder effect social status and subjective well-being. *Psychological Science*, *23*(7), 764–771.

Anderson, M., & Magruder, J. (2012). Learning from the crowd: Regression discontinuity estimates of the effects of an online review database*. *The Economic Journal*, *122*(563), 957–989.

Arnesen, S., & Bergfjord, O. (2014). Prediction markets vs polls – an examination of accuracy for the 2008 and 2012 elections. *The Journal of Prediction Markets*, *8*(3), 24–33.

Arrow, K. J., Forsythe, R., Gorham, M., Hahn, R., Hanson, R., Ledyard, J. O., Levmore, S., Litan, R., Milgrom, P., Nelson, F. D., & others (2008). The promise of prediction markets. *Science*, *320*(5878), 877.

Aumann, R. J. (1976). Agreeing to disagree. *The annals of statistics*, (pp. 1236–

1239).

Avery, C., Resnick, P., & Zeckhauser, R. (1999). The Market for Evaluations. *The American Economic Review*, *89*(3), 564–584.

Bandura, A. (1977). *Social learning theory*, vol. viii. Oxford, England: Prentice-Hall.

Banerjee, A. V. (1992). A Simple Model of Herd Behavior. *The Quarterly Journal of Economics*, *107*(3), 797–817.

Bartling, B., Fehr, E., & Schmidt, K. M. (2012). Screening, competition, and job design: Economic origins of good jobs. *The American Economic Review*, *102*(2), 834–64.

Basu, K. (1994). The traveler's dilemma: Paradoxes of rationality in game theory. *The American Economic Review*, (pp. 391–395).

Bauer, M. A., Wilkie, J. E., Kim, J. K., & Bodenhausen, G. V. (2012). Cuing Consumerism Situational Materialism Undermines Personal and Social Well-Being. *Psychological Science*, *23*(5), 517–523.

Begley, C. G., & Ellis, L. M. (2012). Drug development: Raise standards for preclinical cancer research. *Nature*, *483*(7391), 531–533.

Benaïm, M., Hofbauer, J., & Hopkins, E. (2009). Learning in games with unstable equilibria. *Journal of Economic Theory*, *144*(4), 1694–1709.

Berg, J., Forsythe, R., Nelson, F., & Rietz, T. (2008a). Results from a dozen years of election futures markets research. *Handbook of Experimental Economics Results*, *1*, 742–751.

Berg, J. E., Nelson, F. D., & Rietz, T. A. (2008b). Prediction market accuracy in the long run. *International Journal of Forecasting*, *24*(2), 285–300.

Berg, J. E., & Rietz, T. A. (2003). Prediction Markets as Decision Support Systems. *Information Systems Frontiers*, *5*(1), 79–93.

Bernheim, B. D. (1984). Rationalizable strategic behavior. *Econometrica: Journal of the Econometric Society*, (pp. 1007–1028).

Bikhchandani, S., Hirshleifer, D., & Welch, I. (1992). A theory of fads, fashion, custom, and cultural change as informational cascades. *Journal of political Economy*, *100*(5), 992–1026.

Bohannon, J. (2014). Replication effort provokes praise—and 'bullying' charges. *Science*, *344*(6186), 788–789.

Bohm, P., & Sonnegård, J. (1999). Political stock markets and unreliable polls.

The Scandinavian Journal of Economics, *101*(2), 205–222.

Borghesi, R. (2009). An Examination of Prediction Market Efficiency: Nba Contracts on Tradesports. *The Journal of Prediction Markets*, *3*(2), 63–77.

Brier, G. W. (1950). Verification of forecasts expressed in terms of probability. *Monthly weather review*, *78*(1), 1–3.

Brodeur, A., Lé, M., Sangnier, M., & Zylberberg, Y. (2016). Star wars: the empirics strike back. *American Economic Journal: Applied Economics*, *8*(1), 1–32.

Button, K. S., Ioannidis, J. P., Mokrysz, C., Nosek, B. A., Flint, J., Robinson, E. S., & Munafò, M. R. (2013). Power failure: Why small sample size undermines the reliability of neuroscience. *Nature Reviews Neuroscience*, *14*(5), 365–376.

Cabral, L., & Hortacsu, A. (2010). The dynamics of seller reputation: Evidence from ebay. *The Journal of Industrial Economics*, *58*(1), 54–78.

Cai, H., Chen, Y., & Fang, H. (2009). Observational Learning: Evidence from a Randomized Natural Field Experiment. *The American Economic Review*, *99*(3), 864.

Camerer, C., & Hua Ho, T. (1999). Experience-weighted attraction learning in normal form games. *Econometrica*, *67*(4), 827–874.

Camerer, C. F. (1989). An experimental test of several generalized utility theories. *Journal of Risk and uncertainty*, *2*(1), 61–104.

Camerer, C. F., Dreber, A., Forsell, E., Ho, T.-H., Huber, J., Johannesson, M., Kirchler, M., Almenberg, J., Altmejd, A., Chan, T., Heikensten, E., Holzmeister, F., Imai, T., Isaksson, S., Nave, G., Pfeiffer, T., Razen, M., & Wu, H. (2016). Evaluating replicability of laboratory experiments in economics. *Science*.

Cameron, A. C., & Miller, D. L. (2015). A practitioner's guide to cluster-robust inference. *Journal of Human Resources*, *50*(2), 317–372.

Carpenter, S. (2012). Psychology's bold initiative. *Science*, *335*(6076), 1558–1561.

Casey, K., Glennerster, R., & Miguel, E. (2012). Reshaping Institutions: Evidence on Aid Impacts Using a Preanalysis Plan*. *The Quarterly Journal of Economics*, *127*(4), 1755–1812.

Chan, Y. Y., & Ngai, E. W. T. (2011). Conceptualising electronic word of

mouth activity: An input-process-output perspective. *Marketing Intelligence & Planning*, *29*(5), 488–516.

Charness, G., & Dufwenberg, M. (2011). Participation. *The American Economic Review*, *101*(4), 1211–1237.

Chen, D. L., Schonger, M., & Wickens, C. (2015). oTree-An Open-Source Platform for Laboratory, Online, and Field Experiments. Working Paper.

Chen, R., & Chen, Y. (2011). The potential of social identity for equilibrium selection. *The American Economic Review*, *101*(6), 2562–2589.

Chen, Y. (2005). *Markets as an information aggregation mechanism for decision support*. Doctor of Philosophy Thesis, School of Information Sciences and Technology, The Pennsylvania State University.

Chen, Y., Kash, I. A., Ruberry, M., & Shnayder, V. (2014). Eliciting Predictions and Recommendations for Decision Making. *ACM Transactions on Economics and Computation*, *2*(2).

Chen, Y., Ruberry, M., & Wortman Vaughan, J. (2012). Designing informative securities. In *Proceedings of the 28th Conference on Uncertainty in Artificial Intelligence (UAI)*, (pp. 185–195).

Chen, Y., Wang, Q., & Xie, J. (2011). Online social interactions: A natural experiment on word of mouth versus observational learning. *Journal of marketing research*, *48*(2), 238–254.

Chetty, R., Looney, A., & Kroft, K. (2009). Salience and Taxation: Theory and Evidence. *The American Economic Review*, *99*(4), 1145.

Chevalier, J. A., & Mayzlin, D. (2006). The effect of word of mouth on sales: Online book reviews. *Journal of marketing research*, *43*(3), 345–354.

Cowgill, B., Wolfers, J., & Zitzewitz, E. (2009). Using Prediction Markets to Track Information Flows: Evidence from Google. In *AMMA*, (p. 3).

Cox, J. C., & Oaxaca, R. L. (1995). Inducing risk-neutral preferences: further analysis of the data. *Journal of Risk and Uncertainty*, *11*(1), 65–79.

Critcher, C. R., & Gilovich, T. (2008). Incidental environmental anchors. *Journal of Behavioral Decision Making*, *21*(3), 241–251.

Cumming, G. (2008). Replication and p intervals: p values predict the future only vaguely, but confidence intervals do much better. *Perspectives on Psychological Science*, *3*(4), 286–300.

De Clippel, G., Eliaz, K., & Knight, B. G. (2014). On the selection of arbitra-

tors. *The American Economic Review*, *104*(11), 3434–58.

de Leeuw, J. R. (2014). jsPsych: A JavaScript library for creating behavioral experiments in a Web browser. *Behavior research methods*, *47*(1), 1–12.

Debnath, S., Pennock, D. M., Giles, C. L., & Lawrence, S. (2003). Information incorporation in online in-game sports betting markets. In *Proceedings of the 4th ACM conference on Electronic commerce*, (pp. 258–259). New York, NY, USA: ACM.

Dewald, W. G., Thursby, J. G., & Anderson, R. G. (1986). Replication in empirical economics: The journal of money, credit and banking project. *The American Economic Review*, *76*(4), 587–603.

Doyen, S., Klein, O., Pichon, C.-L., & Cleeremans, A. (2012). Behavioral priming: It's all in the mind, but whose mind? *PLoS ONE*, *7*(1), e29081.

Dreber, A., Pfeiffer, T., Almenberg, J., Isaksson, S., Wilson, B., Chen, Y., Nosek, B. A., & Johannesson, M. (2015). Using prediction markets to estimate the reproducibility of scientific research. *Proceedings of the National Academy of Sciences*, *112*(50), 15343–15347.

Duffy, J., & Puzzello, D. (2014). Gift exchange versus monetary exchange: Theory and evidence. *The American Economic Review*, *104*(6), 1735–1776.

Duflo, E., & Saez, E. (2002). Participation and investment decisions in a retirement plan: The influence of colleagues' choices. *Journal of public Economics*, *85*(1), 121–148.

Dulleck, U., Kerschbamer, R., & Sutter, M. (2011). The economics of credence goods: An experiment on the role of liability, verifiability, reputation, and competition. *The American Economic Review*, *101*(2), 526–55.

Dwan, K., Altman, D. G., Arnaiz, J. A., Bloom, J., Chan, A.-W., Cronin, E., Decullier, E., Easterbrook, P. J., Von Elm, E., Gamble, C., Ghersi, D., Ioannidis, J. P. A., Simes, J., & Williamson, P. R. (2008). Systematic Review of the Empirical Evidence of Study Publication Bias and Outcome Reporting Bias. *PLoS ONE*, *3*(8), e3081.

Eckel, C. C., & Petrie, R. (2011). Face value. *The American Economic Review*, *101*(4), 1497–1513.

Erev, I., & Roth, A. E. (1998). Predicting how people play games: Reinforcement learning in experimental games with unique, mixed strategy equilibria. *American economic review*, (pp. 848–881).

Ericson, K. M. M., & Fuster, A. (2011). Expectations as endowments: Evidence on reference-dependent preferences from exchange and valuation experiments. *The Quarterly Journal of Economics*, *126*(4), 1879–1907.

Erikson, R. S., & Wlezien, C. (2008). Are political markets really superior to polls as election predictors? *Public Opinion Quarterly*, *72*(2), 190–215.

Erikson, R. S., & Wlezien, C. (2012). Markets vs. polls as election predictors: An historical assessment. *Electoral Studies*, *31*(3), 532–539.

Erkal, N., Gangadharan, L., & Nikiforakis, N. (2011). Relative earnings and giving in a real-effort experiment. *The American Economic Review*, *101*(7), 3330–3348.

Fehr, E., Herz, H., & Wilkening, T. (2013). The lure of authority: Motivation and incentive effects of power. *The American Economic Review*, *103*(4), 1325–1359.

Feigenbaum, J., Fortnow, L., Pennock, D. M., & Sami, R. (2005). Computation in a distributed information market. *Theoretical Computer Science*, *343*, 114–132.

Finkelstein, A. (2009). E-ztax: Tax Salience and Tax Rates. *The Quarterly journal of economics*, *124*(3), 969–1010.

Fischbacher, U. (2007). z-Tree: Zurich toolbox for ready-made economic experiments. *Experimental economics*, *10*(2), 171–178.

Forsell, E. (2016). Trading performance in prediction markets with different structures. Working Paper.

Forsell, E., Pfeiffer, T., Almenberg, J., Wilson, B., Chen, Y., Nosek, B. A., Johannesson, M., & Dreber, A. (2016). Using Prediction Markets to Estimate the Reproducibility in Science - The Many Labs 2 replications. Working Paper.

Forsythe, R., Nelson, F., Neumann, G. R., & Wright, J. (1992). Anatomy of an experimental political stock market. *The American Economic Review*, *85*(5), 1142–1161.

Freedman, L. P., Cockburn, I. M., & Simcoe, T. S. (2015). The economics of reproducibility in preclinical research. *PLoS Biol*, *13*(6), e1002165.

Friberg, R., & Grönqvist, E. (2012). Do expert reviews affect the demand for wine? *American Economic Journal: Applied Economics*, *4*(1), 193–211.

Friedman, D., & Oprea, R. (2012). A continuous dilemma. *The American*

Economic Review, *102*(1), 337–363.

Fudenberg, D., Rand, D. G., & Dreber, A. (2012). Slow to anger and fast to forgive: Cooperation in an uncertain world. *The American Economic Review*, *102*(2), 720–749.

Gabaix, X., & Laibson, D. (2006). Shrouded attributes, consumer myopia, and information suppression in competitive markets. *Quarterly Journal of Economics*, *121*(2).

Gelman, A., & Carlin, J. (2014). Beyond Power Calculations Assessing Type S (Sign) and Type M (Magnitude) Errors. *Perspectives on Psychological Science*, *9*(6), 641–651.

Gelman, A., & Loken, E. (2013). The garden of forking paths: Why multiple comparisons can be a problem, even when there is no 'fishing expedition'or 'p-hacking'and the research hypothesis was posited ahead of time. *Downloaded January*, *30*, 2014.

Gelman, A., & Stern, H. (2006). The difference between "significant" and "not significant" is not itself statistically significant. *The American Statistician*, *60*(4), 328–331.

Gerber, A. S., & Malhotra, N. (2008). Publication bias in empirical sociological research: Do arbitrary significance levels distort published results? *Sociological Methods & Research*.

Giessner, S. R., & Schubert, T. W. (2007). High in the hierarchy: How vertical location and judgments of leaders' power are interrelated. *Organizational Behavior and Human Decision Processes*, *104*(1), 30–44.

Gil, R. G. R., & Levitt, S. D. (2007). Testing the efficiency of markets in the 2002 world cup. *The Journal of Prediction Markets*, *1*(3), 255–270.

Gill, D., & Prowse, V. (2012). A structural analysis of disappointment aversion in a real effort competition. *The American Economic Review*, *102*(1), 469–503.

Gimpel, H., & Teschner, F. (2014). Market-Based Collective Intelligence in Enterprise 2.0 Decision Making. SSRN Scholarly Paper ID 2401590, Social Science Research Network, Rochester, NY.

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

Goel, S., Pennock, D., Reeves, D. M., & Yu, C. (2008). Yoopick: A Combinatorial Sports Prediction Market. In *AAAI*, (pp. 1880–1881).

Google (2012). Google Play for Developers: Ratings and Comments.

Graham, J., Haidt, J., & Nosek, B. A. (2009). Liberals and conservatives rely on different sets of moral foundations. *Journal of personality and social psychology*, *96*(5), 1029.

Gray, K., & Wegner, D. M. (2009). Moral typecasting: divergent perceptions of moral agents and moral patients. *Journal of personality and social psychology*, *96*(3), 505.

Hanson, R. (1990). Market-based foresight-A proposal. *Foresight Update*, *10*(1), 3.

Hanson, R. (1992). Idea futures: encouraging an honest consensus. *Extropy*, *3*(2), 7–17.

Hanson, R. (1995). Could gambling save science? Encouraging an honest consensus. *Social Epistemology*, *9*(1), 3–33.

Hanson, R. (1999). Decision markets. *IEEE Intelligent Systems*, *14*(3), 16–19.

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

Hanson, R. (2006). Decision markets for policy advice. In E. M. Patashnik, & A. S. Gerber (Eds.) *Promoting the general welfare: American democracy and the political economy of government performance*, (pp. 151–173). Washington D.C.: Brookings Institution Press.

Hanson, R. (2007). Logarithmic market scoring rules for modular combinatorial information aggregation. *The Journal of Prediction Markets*, *1*(1), 3–15.

Hauser, M., Cushman, F., Young, L., Kang-Xing Jin, R., & Mikhail, J. (2007). A dissociation between moral judgments and justifications. *Mind & language*, *22*(1), 1–21.

Hendriks, A. (2012). SoPHIE—Software Platform for Human Interaction Experiments. Working Paper, Working Paper, University of Osnabrück.

Hesselius, P., Johansson, P., & Nilsson, J. P. (2009). Sick of Your Colleagues' Absence? *Journal of the European Economic Association*, *7*(2/3), 583–594.

Hewitt, J. K. (2012). Editorial policy on candidate gene association and candidate gene-by-environment interaction studies of complex traits. *Behavior genetics*, *42*(1), 1–2.

Hilger, J., Rafert, G., & Villas-Boas, S. (2011). Expert opinion and the demand for experience goods: an experimental approach in the retail wine market. *Review of Economics and Statistics*, *93*(4), 1289–1296.

Hofbauer, J., & Sandholm, W. H. (2007). Evolution in games with randomly disturbed payoffs. *Journal of Economic Theory*, *132*(1), 47–69.

Hopkins, E. (2002). Two competing models of how people learn in games. *Econometrica*, (pp. 2141–2166).

Horn, C. F., Ivens, B. S., Ohneberg, M., & Brem, A. (2014). Prediction markets – a literature review 2014. *The Journal of Prediction Markets*, *8*(2), 89–126.

Hsee, C. K. (1998). Less is better: When low-value options are valued more highly than high-value options. *Journal of Behavioral Decision Making*, *11*.

Huang, Y., Tse, C.-S., & Cho, K. W. (2014). Living in the north is not necessarily favorable: Different metaphoric associations between cardinal direction and valence in Hong Kong and in the United States. *European Journal of Social Psychology*, *44*(4), 360–369.

Huck, S., Seltzer, A. J., & Wallace, B. (2011). Deferred compensation in multiperiod labor contracts: An experimental test of Lazear's model. *The American Economic Review*, *101*(2), 819–843.

Humphreys, M., Sierra, R. S. d. l., & Windt, P. v. d. (2013). Fishing, Commitment, and Communication: A Proposal for Comprehensive Nonbinding Research Registration. *Political Analysis*, *21*(1), 1–20.

Ifcher, J., & Zarghamee, H. (2011). Happiness and time preference: The effect of positive affect in a random-assignment experiment. *The American Economic Review*, *101*(7), 3109–3129.

Imbens, G. W., & Lemieux, T. (2008). Regression discontinuity designs: A guide to practice. *Journal of econometrics*, *142*(2), 615–635.

Inbar, Y., Pizarro, D. A., Knobe, J., & Bloom, P. (2009). Disgust sensitivity predicts intuitive disapproval of gays. *Emotion*, *9*(3), 435.

Ioannidis, J., & Doucouliagos, C. (2013). What's to know about the credibility of empirical economics? *Journal of Economic Surveys*, *27*(5), 997–1004.

Ioannidis, J. P., Munafo, M. R., Fusar-Poli, P., Nosek, B. A., & David, S. P. (2014). Publication and other reporting biases in cognitive sciences: detection, prevalence, and prevention. *Trends in cognitive sciences*, *18*(5), 235–241.

Ioannidis, J. P. A. (2005). Why most published research findings are false. *PLoS*

Med, *2*(8), e124.

Ioannidis, J. P. A. (2008). Why most discovered true associations are inflated. *Epidemiology*, *19*(5), 640–648.

Ioannidis, J. P. A. (2014). How to Make More Published Research True. *PLoS Med*, *11*(10), e1001747.

Iyer, K., Johari, R., & Moallemi, C. C. (2010). Information aggregation in smooth markets. In *Proceedings of the 11th ACM conference on Electronic commerce*, (pp. 199–206). ACM.

Jin, G. Z., & Kato, A. (2006). Price, quality, and reputation: Evidence from an online field experiment. *The RAND Journal of Economics*, *37*(4), 983–1005.

John, L. K., Loewenstein, G., & Prelec, D. (2012). Measuring the Prevalence of Questionable Research Practices With Incentives for Truth Telling. *Psychological Science*, *23*(5), 524–532.

Kahneman, D. (2012). A proposal to deal with questions about priming effects.

Kahneman, D., & Tversky, A. (1979). Prospect theory: An analysis of decision under risk. *Econometrica: Journal of the Econometric Society*, (pp. 263–291).

Katz, M. L., & Shapiro, C. (1985). Network Externalities, Competition, and Compatibility. *The American Economic Review*, *75*(3), 424–440.

Kay, A. C., Laurin, K., Fitzsimons, G. M., & Landau, M. J. (2014). A functional basis for structure-seeking: Exposure to structure promotes willingness to engage in motivated action. *Journal of Experimental Psychology: General*, *143*(2), 486.

Kessler, J. B., & Roth, A. E. (2012). Organ allocation policy and the decision to donate. *The American Economic Review*, *102*(5), 2018–2047.

Kirchler, M., Huber, J., & Stöckl, T. (2012). Thar she bursts: Reducing confusion reduces bubbles. *The American Economic Review*, *102*(2), 865–883.

Klein, R., Nosek, B. A., Vianello, M., & Hasselman, F. (2014). Many Labs 2: Investigating Variation in Replicability Across Sample and Setting. Proposal.

Klein, R. A., Ratliff, K. A., Vianello, M., Adams Jr, R. B., Bernstein, M. J., Bocian, K., Brandt, M. J., Brooks, B., Brumbaugh, C. C., & others (2015a). Investigating variation in replicability. *Social Psychology*.

Klein, R. A., Ratliff, K. A., Vianello, M., Adams Jr, R. B., Bernstein, M. J., Bocian, K., Brandt, M. J., Brooks, B., Brumbaugh, C. C., & others (2015b). Many Labs 3: Evaluating participant pool quality across the academic

semester via replication. Working Paper.

Knobe, J. (2003). Intentional action and side effects in ordinary language. *Analysis*, *63*(279), 190–194.

Kogan, S., Kwasnica, A. M., & Weber, R. A. (2011). Coordination in the presence of asset markets. *The American Economic Review*, *101*(2), 927–47.

Kuziemko, I., Buell, R. W., Reich, T., & Norton, M. I. (2014). "Last-place aversion": Evidence and redistributive implications. *The Quarterly Journal of Economics*, *129*(1), 105–149.

Lacetera, N., Pope, D. G., & Sydnor, J. R. (2012). Heuristic Thinking and Limited Attention in the Car Market. *American Economic Review*, *102*(5), 2206–2236.

Lane, D. M., & Dunlap, W. P. (1978). Estimating effect size: Bias resulting from the significance criterion in editorial decisions. *British Journal of Mathematical and Statistical Psychology*, *31*(2), 107–112.

Lawrence, M. S., Stojanov, P., Polak, P., Kryukov, G. V., Cibulskis, K., Sivachenko, A., Carter, S. L., Stewart, C., Mermel, C. H., Roberts, S. A., & others (2013). Mutational heterogeneity in cancer and the search for new cancer-associated genes. *Nature*, *499*(7457), 214–218.

Leamer, E. E. (1983). Let's take the con out of econometrics. *The American Economic Review*, *73*(1), 31–43.

Ledyard, J., Hanson, R., & Ishikida, T. (2009). An experimental test of combinatorial information markets. *Journal of Economic Behavior & Organization*, *69*(2), 182–189.

Lee, D. S., & Lemieux, T. (2010). Regression Discontinuity Designs in Economics. *Journal of Economic Literature*, *48*(2), 281–355.

Leek, J. T., Patil, P., & Peng, R. D. (2015). A glass half full interpretation of the replicability of psychological science. *arXiv:1509.08968*.

Lewis, D. (1969). *Convention: a philosophical study*. Cambridge, Harvard university press.

Li, X., & Wu, L. (2013). Measuring effects of observational learning and social-network word-of-mouth (WOM) on the sales of daily-deal vouchers. In *System Sciences (HICSS), 2013 46th Hawaii International Conference on*, (pp. 2908–2917). IEEE.

Luca, M. (2011). Reviews, Reputation, and Revenue: The Case of Yelp.com.

Working Paper No. 12-016. (Revise and resubmit at the American Economic Journal - Applied Economics.).

Lucking-Reiley, D., Bryan, D., Prasad, N., & Reeves, D. (2007). Pennies from ebay: The determinants of price in online auctions. *The Journal of Industrial Economics*, *55*(2), 223–233.

Maniadis, Z., Tufano, F., & List, J. A. (2014). One swallow doesn't make a summer: New evidence on anchoring effects. *The American Economic Review*, *104*(1), 277–90.

Manski, C. F. (2006). Interpreting the predictions of prediction markets. *Economics Letters*, *91*(3), 425–429.

Masicampo, E. J., & Lalande, D. R. (2012). A peculiar prevalence of p values just below .05. *The Quarterly Journal of Experimental Psychology*, *65*(11), 2271–2279.

McCarthy, J. (1956). Measures of the value of information. *Proceedings of the National Academy of Sciences of the United States of America*, *42*(9), 654.

McCrary, J. (2008). Manipulation of the running variable in the regression discontinuity design: A density test. *Journal of Econometrics*, *142*(2), 698–714.

McCullough, B. D., McGeary, K. A., & Harrison, T. D. (2006). Lessons from the JMCB Archive. *Journal of Money, Credit, and Banking*, *38*(4), 1093–1107.

McCullough, B. D., & Vinod, H. D. (2003). Verifying the solution from a nonlinear solver: A case study. *The American Economic Review*, *93*, 873–892.

McFadden, D. (1974). Conditional logit analysis of qualitative choice behavior. In P. Zarembka (Ed.) *Frontiers of Econometrics*. Academic Press.

McNutt, M. (2014). Reproducibility. *Science*, *343*(6168), 229–229.

Melnik, M. I., & Alm, J. (2002). Does a Seller's Ecommerce Reputation Matter? Evidence from eBay Auctions. *The Journal of Industrial Economics*, *50*(3), 337–349.

Merlob, B., Plott, C. R., & Zhang, Y. (2012). The CMS auction: Experimental studies of a median-bid procurement auction with nonbinding bids. *The Quarterly Journal of Economics*, *127*(2), 793–827.

Miguel, E., Camerer, C., Casey, K., Cohen, J., Esterling, K. M., Gerber, A., Glennerster, R., Green, D. P., Humphreys, M., Imbens, G., Laitin, D.,

Madon, T., Nelson, L., Nosek, B. A., Petersen, M., Sedlmayr, R., Simmons, J. P., Simonsohn, U., & Van der Laan, M. (2014). Promoting Transparency in Social Science Research. *Science (New York, N.Y.)*, *343*(6166), 30–31.

Miyamoto, Y., & Kitayama, S. (2002). Cultural variation in correspondence bias: the critical role of attitude diagnosticity of socially constrained behavior. *Journal of personality and social psychology*, *83*(5), 1239.

Molleman, L., Arechar, A. A., & Gaetcher, S. (2015). LIONESS | Live Online Experiments' Server Software.

Moretti, E. (2011). Social learning and peer effects in consumption: Evidence from movie sales. *The Review of Economic Studies*, *78*(1), 356–393.

Myerson, R., & Weibull, J. (2015). Tenable Strategy Blocks and Settled Equilibria. *Econometrica*, *83*(3), 943–976.

Nash, J. (1950). *Non-cooperative games*. Ph. D. Thesis, Princeton University.

Norenzayan, A., Smith, E. E., Kim, B. J., & Nisbett, R. E. (2002). Cultural preferences for formal versus intuitive reasoning. *Cognitive Science*, *26*(5), 653–684.

Nosek, B. A., Alter, G., Banks, G. C., Borsboom, D., Bowman, S. D., Breckler, S. J., Buck, S., Chambers, C. D., Chin, G., Christensen, G., & others (2015). Promoting an open research culture: Author guidelines for journals could help to promote transparency, openness, and reproducibility. *Science*, *348*(6242), 1422.

Nuzzo, R. (2014). Statistical errors. *Nature*, *506*(7487), 150–152.

Open Science Collaboration (2012). An open, large-scale, collaborative effort to estimate the reproducibility of psychological science. *Perspectives on Psychological Science*, *7*(6), 657–660.

Open Science Collaboration (2015). Estimating the reproducibility of psychological science. *Science*, *349*(6251).

Ostrovsky, M. (2012). Information aggregation in dynamic markets with strategic traders. *Econometrica*, *80*(6), 2595–2647.

Othman, A., & Sandholm, T. (2010). Decision rules and decision markets. In *Proceedings of the 9th International Conference on Autonomous Agents and Multiagent Systems: volume 1-Volume 1*, (pp. 625–632). International Foundation for Autonomous Agents and Multiagent Systems.

O'Leary, D. E. (2011). Prediction markets as a forecasting tool. *Advances in*

business and management forecasting, *8*, 169–184.

Page, L. (2008). Comparing prediction market prices and opinion polls in political elections. *The Journal of Prediction Markets*, *2*(1), 91–97.

Palan, S. (2015). GIMS—Software for asset market experiments. *Journal of Behavioral and Experimental Finance*, *5*, 1–14.

Pashler, H., & Harris, C. R. (2012). Is the Replicability Crisis Overblown? Three Arguments Examined. *Perspectives on Psychological Science*, *7*(6), 531–536.

Paul, R. J., & Weinbach, A. P. (2009). Sportsbook Behavior in the Ncaa Football Betting Market: Tests of the Traditional and Levitt Models of Sportsbook Behavior. *The Journal of Prediction Markets*, *3*(2), 21–37.

Pearce, D. G. (1984). Rationalizable strategic behavior and the problem of perfection. *Econometrica: Journal of the Econometric Society*, (pp. 1029–1050).

Pettit, J., Friedman, D., Kephart, C., & Oprea, R. (2014). Software for continuous game experiments. *Experimental Economics*, *17*(4), 631–648.

Plott, C. R., & Chen, K.-Y. (2002). Information aggregation mechanisms: Concept, design and implementation for a sales forecasting problem. Working Paper No. 1131, Division of the Humanitites and Social Sciences, California Institute of Technology, Pasadena.

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

Pope, D. G. (2009). Reacting to rankings: evidence from "America's Best Hospitals". *Journal of health economics*, *28*(6), 1154–1165.

Prinz, F., Schlange, T., & Asadullah, K. (2011). Believe it or not: How much can we rely on published data on potential drug targets? *Nature Reviews Drug Discovery*, *10*(9), 712–712.

Rabin, M. (2000). Risk Aversion and Expected-Utility Theory: A Calibration Theorem. *Econometrica*, *68*(5), 1281–1292.

Reinstein, D. A., & Snyder, C. M. (2005). The influence of expert reviews on consumer demand for experience goods: A case study of movie critics*. *The journal of industrial economics*, *53*(1), 27–51.

Resnick, P., Zeckhauser, R., Swanson, J., & Lockwood, K. (2006). The value of reputation on eBay: A controlled experiment. *Experimental Economics*,

9(2), 79–101.

Rhode, P. W. (2009). The emergence of prediction markets within business firms: a skeptical perspective from an intrigued academic. *The Journal of Prediction Markets*, *3*(1), 87–88.

Rieg, R., & Schoder, R. (2010). Forecasting Accuracy: Comparing Prediction Markets and Surveys - an Experimental Study. *The Journal of Prediction Markets*, *4*(3), 1–19.

Risen, J. L., & Gilovich, T. (2008). Why people are reluctant to tempt fate. *Journal of personality and social psychology*, *95*(2), 293.

Ritchie, S. J., Wiseman, R., & French, C. C. (2012). Failing the future: Three unsuccessful attempts to replicate Bem's 'retroactive facilitation of recall' effect. *PLoS ONE*, *7*(3), e33423.

Ross, L., Greene, D., & House, P. (1977). The "false consensus effect": An egocentric bias in social perception and attribution processes. *Journal of experimental social psychology*, *13*(3), 279–301.

Roth, A. E. (1994). Lets keep the con out of experimental econ.: A methodological note. *Empirical Economics*, *19*(2), 279–89.

Roth, A. E., & Erev, I. (1995). Learning in extensive-form games: Experimental data and simple dynamic models in the intermediate term. *Games and economic behavior*, *8*(1), 164–212.

Roth, A. E., & Malouf, M. W. (1979). Game-theoretic models and the role of information in bargaining. *Psychological review*, *86*(6), 574.

Rothschild, D. (2009). Forecasting elections comparing prediction markets, polls, and their biases. *Public Opinion Quarterly*, *73*(5), 895–916.

Rottenstreich, Y., & Hsee, C. K. (2001). Money, kisses, and electric shocks: On the affective psychology of risk. *Psychological Science*, *12*(3), 185–190.

Salganik, M. J., Dodds, P. S., & Watts, D. J. (2006). Experimental study of inequality and unpredictability in an artificial cultural market. *science*, *311*(5762), 854–856.

Sandholm, W. H. (2010). *Population games and evolutionary dynamics*. MIT press.

Savage, L. J. (1954). *The foundations of statistics*. New York: Wiley.

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

Savani, K., Markus, H. R., Naidu, N. V. R., Kumar, S., & Berlia, N. (2010). What counts as a choice? US Americans are more likely than Indians to construe actions as choices. *Psychological Science*, *21*(3), 391–398.

Schwarz, N., Strack, F., & Mai, H.-P. (1991). Assimilation and contrast effects in part-whole question sequences: A conversational logic analysis. *Public opinion quarterly*, *55*(1), 3–23.

Selten, R., Sadrieh, A., & Abbink, K. (1999). Money does not induce risk neutral behavior, but binary lotteries do even worse. *Theory and Decision*, *46*(3), 213–252.

Shafir, E. (1993). Choosing versus rejecting: Why some options are both better and worse than others. *Memory & cognition*, *21*(4), 546–556.

Simmons, J. P., Nelson, L. D., & Simonsohn, U. (2011). False-positive psychology undisclosed flexibility in data collection and analysis allows presenting anything as significant. *Psychological Science*, *22*(11), 1359–1366.

Simon, H. A. (1955). A Behavioral Model of Rational Choice. *The Quarterly Journal of Economics*, *69*(1), 99–118.

Simonsohn, U. (2015). Small telescopes detectability and the evaluation of replication results. *Psychological Science*, *26*(5), 559–569.

Simonsohn, U., Nelson, L. D., & Simmons, J. P. (2013). P-Curve: A Key to the File Drawer. *Journal of Experimental Psychology: General, Forthcoming*.

Smith, C. A. (1961). Consistency in statistical inference and decision. *Journal of the Royal Statistical Society. Series B (Methodological)*, (pp. 1–37).

Smith, V. L. (1962). An Experimental Study of Competitive Market Behavior. *Journal of Political Economy*, *70*(2), 111–137.

Sobel, R. S., & Raines, S. T. (2003). An examination of the empirical derivatives of the favourite-longshot bias in racetrack betting. *Applied Economics*, *35*(4), 371–385.

Sonnemann, U., Camerer, C. F., Fox, C. R., & Langer, T. (2013). How psychological framing affects economic market prices in the lab and field. *PNAS*, *110*(29), 11779–11784.

Sorensen, A. T. (2006). Social Learning and Health Plan Choice. *The RAND Journal of Economics*, *37*(4), 929–945.

Stroebe, W., Postmes, T., & Spears, R. (2012). Scientific Misconduct and the Myth of Self-Correction in Science. *Perspectives on Psychological Science*,

7(6), 670–688.

Thaler, R. H., & Ziemba, W. T. (1988). Parimutuel betting markets: Racetracks and lotteries. *Journal of Economic perspectives*, *2*(2), 161–174.

Tucker, C., & Zhang, J. (2011). How does popularity information affect choices? A field experiment. *Management Science*, *57*(5), 828–842.

Tversky, A., & Gati, I. (1978). Studies of similarity. *Cognition and categorization*, *1*(1978), 79–98.

Tversky, A., & Kahneman, D. (1981). The framing of decisions and the psychology of choice. *Science*, *211*(4481), 453–458.

Tziralis, G., & Tatsiopoulos, I. (2007). Prediction markets: An extended literature review. *The Journal of Prediction Markets*, *1*(1), 75–91.

Van Lange, P. A., De Bruin, E., Otten, W., & Joireman, J. A. (1997). Development of prosocial, individualistic, and competitive orientations: theory and preliminary evidence. *Journal of personality and social psychology*, *73*(4), 733.

Verhagen, J., & Wagenmakers, E.-J. (2014). Bayesian tests to quantify the result of a replication attempt. *Journal of Experimental Psychology: General*, *143*(4), 1457.

Von Neumann, J., & Morgenstern, O. (1947). *Theory of games and economic behavior, 2nd rev. ed.*. Princeton, NJ, US: Princeton university press, 2nd rev. ed.

Wolfers, J., & Leigh, A. (2002). Three tools for forecasting federal elections: lessons from 2001. *Australian Journal of Political Science*, *37*(2), 223–240.

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

Wolfers, J., & Zitzewitz, E. (2006). Interpreting prediction market prices as probabilities. Working Paper No. 12200, National Bureau of Economic Research.

Young, H. P. (1993). The evolution of conventions. *Econometrica: Journal of the Econometric Society*, (pp. 57–84).

Zaval, L., Keenan, E. A., Johnson, E. J., & Weber, E. U. (2014). How warm days increase belief in global warming. *Nature Climate Change*, *4*(2), 143–147.

Zhong, C.-B., & Liljenquist, K. (2006). Washing away your sins: Threatened morality and physical cleansing. *Science*, *313*(5792), 1451–1452.