Difficult Choices

Essays on Economic Behavior

Johan Almenberg

Dissertation for the Degree of Doctor of Philosophy
Stockholm School of Economics

KEYWORDS: behavioral economics, financial literacy, housing market, transmission of inequality, marriage, status, prediction markets, price-quality heuristic, social preferences

ISBN NR 978-91-7258-792-2

PRINTED BY:
Elanders, Vällingby 2009

DISTRIBUTED BY:
EFI, The Economic Research Institute
Stockholm School of Economics
P O Box 6501, SE-113 83 Stockholm
www.hhs.se/efi

Till Gustaf, Sandra, Erik och Anna

EFI THE ECONOMIC RESEARCH INSTITUTE

EFI Mission

EFI, the Economic Research Institute at the Stockholm School of Economics, is a scientific institution that works independently of economic, political and sectional interests. It conducts theoretical and empirical research in the management and economic sciences, including selected related disciplines. The Institute encourages and assists in the publication and distribution of its research findings and is also involved in the doctoral education at the Stockholm School of Economics. At EFI, the researchers select their projects based on the need for theoretical or practical development of a research domain, on their methodological interests, and on the generality of a problem.

Research Organization

The research activities at the Institute are organized into 20 Research Centres.
Centre Directors are professors at the Stockholm School of Economics.

EFI Research Centre:	*Centre Director:*
Management and Organisation (A)	Sven-Erik Sjöstrand
Centre for Entrepreneurship and Business Creation (E)	Carin Holmquist
Public Management (F)	Nils Brunsson
Information Management (I)	Mats Lundeberg
Centre for People and Organization (PMO)	Andreas Werr
Centre for Innovation and Operations Management (T)	Pär Åhlström
Centre for Media and Economic Psychology (P)	Richard Wahlund
Centre for Consumer Marketing (CCM)	Magnus Söderlund
Centre for Information and Communication Research (CIC)	Per Andersson
Marketing, Distribution and Industrial Dynamics (D)	Björn Axelsson
Centre for Strategy and Competitiveness (CSC)	Örjan Sölvell
Accounting and Managerial Finance (B)	Johnny Lind
Centre for Financial Analysis and Managerial Economics in Accounting (BFAC)	Kenth Skogsvik
Finance (FI)	Clas Bergström
Centre for Health Economics (CHE)	Magnus Johannesson
International Economics and Geography (IEG)	Mats Lundahl
Economics (S)	Paul Segerstrom
Economic Statistics (ES)	Anders Westlund
Centre for Business Law (RV)	Johnny Herre
Centre for Tax Law (SR)	Bertil Wiman

Chair of the Board: Professor Carin Holmquist
Director: Associate Professor Filip Wijkström

Address

EFI, Box 6501, SE-113 83 Stockholm, Sweden • Website: www.hhs.se/efi/
Telephone: +46(0)8-736 90 00 • Fax: +46(0)8-31 62 70 • E-mail efi@hhs.se

Contents

Acknowledgments

It has been my privilege to attend the Ph.D. program at the Stockholm School of Economics. I would like express my gratitude to the Department of Economics – to my fellow Ph.D. students, to my advisor Magnus Johannesson and the rest of the faculty, and to the administrative staff – for making it the education that is has been, in the intellectual realm as well as at the personal level. Of course, there's no such thing as a free Ph.D, and I am grateful to the Jacob Wallenberg Foundation, the Jan Wallander and Tom Hedelius Foundation, and the Torsten and Ragnar Söderberg Foundations for financial support.

The theme of this thesis is that we face difficult choices in our lives. For me, becoming an economist was a difficult choice, but one which I have never regretted. I am grateful to Rutger Palmstierna for inspiring me to study economics in the first place, and to Kristian Hartelius, Pontus Hansson, and Andreas Bergh for encouraging me to continue with studies at the graduate level.

During my time as a Ph.D. student I had the pleasure of spending some time at the Initiative for Policy Dialogue in New York, the European Bank for Reconstruction and Development in London, the NBER in Cambridge, and the Federal Reserve Bank in Boston. I would like to thank Shana Hofstadter, Shari Spiegel, Erik Berglöf, Christiane Nickel, Dan Feenberg, Karen Norberg, Lorenz Goette, and Stephan Meier for those rewarding experiences. While in Boston I also benefitted from the wisdom and generosity of Yannis Ioannides. For this I owe thanks not only to Yannis but also to Peter Englund at the Stockholm School of Economics.

I would like to thank my friends and my family for their encouragement and enduring support. In addition, it has been my good fortune to work together with some wonderful and very talented people during these years. I married one of my co-authors; to the others I merely offer my profound gratitude. As for you, Anna, words are not enough to express my appreciation. The Ph.D. program has entailed costs as well as benefits, but the simple fact that it brought us together is sufficient to have made it worthwhile a million times over.

Stockholm, May 2009
Johan Almenberg

Chapters

CHAPTER 1

Introduction

The six papers in this volume cover a wide range of subject matters, including real estate, marriage, betting on scientific opinions, wine, and moral judgment. The common denominator is that the papers are all inspired by the many difficult choices we face in our lives.

Some choices, such as the debt-financed acquisition of a home, are of a distinctly economic nature. Yet, the process by which we make such choices does not always follow a strictly economic logic. This is the perspective taken in the first paper.

Other choices, such as marriage, scientific opinion, which wine to buy, or aspects of moral reasoning, are such that we do not typically think of them as belonging to the domain of economics. But economic analysis can yield important insights also in these areas of life. This is the perspective taken in the five subsequent papers.

A short summary of each paper follows.

"Mental Accounting in the Housing Market"

(with A. Karapetyan)

People make mistakes. To an economist, some mistakes are more interesting than others. For example, systematic mistakes lend themselves well to economic modeling, because they represent a sort of predictable irrationality. Sometimes, markets correct for systematic mistakes. At other times, such as when there are constraints on arbitrage, markets may fail to do so. In this case individual irrationality is reflected at the aggregate level, in the form of equilibrium prices that are wrong from an economic perspective. This paper represents an attempt to trace one such systematic mistake from the individual level to the aggregate level.

We approach the problem at hand from several angles. To begin with, we use a survey to identify a consumer bias with regard to different sources of debt-financing. The survey was directed at Swedish apartment owners. In Sweden, owner-occupied apartments are almost exclusively organized as co-ops. These co-ops can, and often do, take on debt. The degree of leverage differs from case to case. Thus a person who is choosing between different apartments is in effect also chosing between different capital structures. We conjecture that for many individuals co-op debt has the advantage,

relative to individually held debt, of being less salient. By this we mean that the debt is more easily ignored.

The increasingly recognized problem of inadequate financial literacy is likely to exacerbate biases, because individuals fail to perceive the economic costs of such "mental accounting". We discuss recent research on salience and on financial literacy at some length. Both concepts have a wide applicability, and are of interest also to readers who care less about the housing market.

We find that most survey respondents are considerably more well-informed about their individually held debt than about the debt held by their co-op. Strikingly, only a small fraction of respondents had ever considered the possibility of exchanging one form of debt for the other. These findings are highly consistent with our conjecture about a biased view of the co-op capital structure.

To explore how an individual bias might affect the market equilibrium, we construct a model of a market with two types. Naive agents are biased in favour of less salient co-op debt. Sophisticated agent are not biased at all. Aside from this, both agents solve the same optimization problem. An important feature of the market in question is that every apartments is unique. It is not possible to turn on a tap and say "I want exactly the apartment that I have right now and then a little bit more of the same". In our model, it is this heterogeneity that prevents a full separation in the market. Agents of both types allow for the possibility that the apartment of their dreams will have the "wrong" capital structure.

Our model predicts that the market will respond only moderately to a change in the relative price of the two sources of debt. We use a large sample of apartment sales data immediately before and after a recent reform to test this prediction. Our empirical results suggest that the market underreacted to the reform as compared to a fully rational benchmark. This is consistent with the prediction of our model. We conclude the paper by discussing the challenge of trying to use regulation to address economic biases.

"Lady and the Trump: Status and Wealth in the Marriage Market"

(with A. Dreber)

If status and wealth are complements, one might expect individuals with high status but low wealth to seek spouses with high wealth but low status. But how can we measure status? In this paper, we make use of an easily observed indicator of high status, a noble surname, to test the hypothesis that individuals with such surnames have a higher probability of marrying "up" (hypergamy) in terms of wealth.

If wealth begets wealth, dynasties may endure even in otherwise meritocratic societies. And if an indicator of past wealth begets wealth, this too will reinforce the persistence of economic standing across generations. A number of previous studies that have analyzed the intergenerational transmission of economic standing. The mechanisms involved can range from the purely biological, such as a heritable component of IQ, to the strictly environmental, such as social norms or knowledge gained at the breakfast table as a child. This inheritance, in turn, can be counteracted, or reinforced, by the institutions of a society.

As an institution, nobility is an anachronism. Status, however, plays an important role in most societies, making individuals allocate valuable resources to status-enhancing activities. To the extent that nobility is a status marker we would expect it to have an economic value. Swedish law, however, does not permit nobility to be traded in itself. Given this, and bearing in mind that the monarch no longer has the right to create new nobility, the marriage market constitutes the sole remaining conduit for observing the valuation of this asset.

We use a sample of about 200,000 Swedish marriages to test the hypothesis that individuals with a noble surname have a higher probability of marrying into a higher wealth bracket. We control for age, income, education and the individual's own wealth.

Our main finding is a significantly higher probability of hypergamy for members of the nobility. This "nobility premium" is sizeable. The effect is statistically significant and robust to a number of different measures of hypergamy. This finding has implications for the intergenerational transmission of inequality, and for the longevity of the institution of nobility itself.

"Prediction Markets in Science: an Experiment"

(with T. Pfeiffer and K. Kittlitz)

A prediction market is a marketplace for contracts whose payoffs depend on the outcome of a future event. In a well-functioning market, contract prices can be interpreted as forecasts about the outcome of the event, derived from the beliefs of all market participants. Prediction markets can be powerful forecasting tools. They have the potential to aggregate private information, to generate and disseminate a consensus among the market participants, and to provide incentives for information acquisition.

Making reliable predictions is a key objective in science. Thus prediction markets offer potential benefits to scientific research. The dissemination and aggregation properties might be beneficial because knowledge in scientific research is often highly decentralized, resulting in diverging opinions within the research community. Prediction markets can generate a consensus estimate on a scientific question and communicate

it to market participants as well as to outside parties such as funding agencies and policy makers. The consensus disseminated by a functioning market has the potential to be more precise than a consensus obtained from traditional methods such as performing a meta-analysis of more or less biased data from the literature, or sampling expert opinion. A reliable consensus is important in order to generate agreement on what the most important open questions are and thus helps to allocate resources in an optimal way. Moreover, prediction markets can be used to fund research, because scientists that invest in well-designed research projects gain an information advantage that allows them to recoup some of the costs.

We designed an experiment in order to examines the compatibility of prediction markets with the current practice of scientific publication. Participants traded claims representing a set of mutually exclusive scientific hypothesis using a conventional prediction market platform. During the course of the experiment, information was released to the participants, in the form of error-prone test results regarding the hypotheses at hand. We explored three settings. In the first setting, different pieces of information were disclosed to the public, i.e., to all traders simultaneously. In the second setting, participants received private information. In the third setting, each piece of information was private at first, but was subsequently disclosed to the public. An automated, subsidizing market maker provided additional incentives for trading and mitigated liquidity problems.

We find that the third setting combines the advantages of the first and second settings. Market performance was as good as in the setting with public information, and better than in the setting with private information. In contrast to the first setting, participants could benefit from information advantages. In other words, the publication of information does not detract from the functionality of prediction markets. We conclude that for integrating prediction markets into the practice of scientific research it is advantageous to use subsidizing market makers, and to keep markets aligned with current publication practice.

"Do More Expensive Wines Taste Better? Evidence from a Large Sample of Blind Tastings"

(with R. Goldstein, A. Dreber, J.W. Emerson, A. Herschkowitsch, and J. Katz)

Suppose good X appears similar to good Y, but has a considerably higher price. Most of us would probably expect good X to be slightly better than good Y along some dimension. Such an expectation would not imply that we would always purchase good X: the difference in how much we appreciate the two goods may be small, and not

worth the additional expense. The point is simply that it seems reasonable to expect there to be some noticeable additional quality to the more expensive good.

Does wine live up to such an expectation? Perhaps not. We analyze a sample of over 6000 blind tasting observations involving about 500 different tasters and as many wines, and find that the correlation between price and overall rating is in fact small and *negative*, suggesting that individuals on average enjoy more expensive wines slightly *less*.

There is no reason to expect individuals with a great deal of experience of drinking wine to be similar to individuals with little experience in this regard. Using wine training, such as participation in a sommelier course, as a proxy for wine expertise, we were able to separate these "experts" from the rest of the tasters in the sample. We find indications of a small positive correlation for experts. This correlation is only marginally significant (the p-value is just below 0.10, despite the large sample size) and the coefficient is small, suggesting that even for expert tasters the correlation between price and subjective appreciation is rather weak.

In both cases, the coefficients are of a moderate magnitude, but non-negligible, given that wine prices cover such a large range. Suppose that Wine X costs ten times more then Wine Y in dollar terms. In terms of a 100-point scale, such as that used by *Wine Spectator*, our estimates predict that non-experts will assign an overall rating that is four points lower for wine X, whereas experts will assign an overall rating that is seven points higher.

"When Does the Price Affect the Taste? Results from a Wine Experiment"

(with A. Dreber)

The first wine paper looked at how tasters rated wines when they were unaware of the price. As mentioned above, we found that cheap wines get as high, or even higher, ratings as expensive wines. But wines are not typically consumed in blind tastings. This raises the question of what effect knowledge about the price has on ratings. Drawing on an experimental design from consumer research, we designed an experiment that examines how knowledge about the price of a wine, and the time at which the information is received, affects how the wine is perceived. The price tag is a form of extrinsic information, and previous studies have shown that the timing of such extrinsic information makes an important difference: if the consumer gets to experience the product first-hand before being informed about some attribute that is not directly observable, the initial impression has been found to dominate. By contrast,

if consumers are first told about the attribute, then this can generate expectations that significantly affect how the good in question is perceived.

In our experiment, subjects tasted a wine that was either very cheap or very expensive, and then assigned a rating to it. The participants were randomized to three settings. In the first setting, they were not informed about the price of the wine until after the experiment. In the second setting, they were informed about the price before tasting and rating the wine. In the third setting, they were informed about the price before rating the wine but after having tasted it.

Our results suggest that hosts offering wine to their guests can safely reveal the price. Much is gained if the wine is expensive, and little is lost if it is cheap. Disclosing the high price before tasting the wine produces considerably higher ratings, although only from women. Disclosing the low price, by contrast, does not result in lower ratings. Consistent with previous research on the timing of attribute information we find that information provided after the consumer has had first-hand experience of the good has little effect.

Our finding indicates that price not only serves to clear markets, it also serves as a marketing tool. The price tag itself influences expectations, which in turn shape the consumer's experience. Our results also suggest that men and women respond differently to attribute information.

"Third Party Intervention, Group Size and Public Goods"

(with A. Dreber, D. G. Rand, and C. Apicella)

Shared beliefs of what constitutes appropriate behaviour greatly affect human decision making in many social domains, ranging from dress codes and marriage practices to the resolution of personal conflicts and the design of public policies. It has been proposed that salient norms can be identified through examining the use of costly punishment by bystanders, or third parties, in experimental games. Third party punishment involves situations where I punish you, at a cost to myself, for what you have done to somebody else. In the type of non-repeated anonymous interactions that have been explored, third parties can never benefit materially from punishing. Nonetheless, participants in such experiments frequently engage in costly punishment.

Third party punishment has been suggested as an effective tool for enforcing norms and promoting cooperation. It has been shown that third party punishment is more common in larger and more complex societies than in small-scale societies Meanwhile, larger and more complex societies give rise to many situations where one individual can either extract a benefit at a cost to many others or incur a cost in order to benefit many others. Yet, to the best of our knowledge, how third party response to an action

varies according to the number of individuals affected has not been investigated. That is the aim of this paper.

Experiments on this topic typically involve a third party observer who can pay to decrease the payoff of one player who has behaved selfishly (or generously) toward another. We investigate whether third parties are sensitive to the number of players affected by this selfish or generous action. The money transfers in our experiment have the non-rivalrous character of a public good. We allowed dictators to be selfish, fair, or generous, and unlike in previous, related experiments, third parties could not only punish but also reward. Across all variations, responses followed a consistent and intuitive pattern: selfish behavior was punished while generous behavior was rewarded.

Third party response was more pronounced when the dictator transfer had the nonrivalrous character of a public good, in the sense that both (a) the number of recipients increased and (b) the dictator's transfer was multiplied by a constant factor, so that the larger number of recipients did not reduce potential payoff of each recipient. Third party response did not change significantly when either of these manipulations was performed alone, suggesting a specific response to situations involving a public good.

CHAPTER 2

Mental Accounting in the Housing Market

with Artashes Karapetyan[1]

Abstract. We use a survey to identify a consumer bias with regard to different sources of debt-financing. Less salient debt may generate psychological benefits. This should be weighed against the possible economic costs of a sub-optimal capital structure, but low levels of financial literacy make it unlikely that all households perceive the full economic costs. As a result there is a bias in favour of less salient debt. In a market with limited scope for arbitrage this consumer bias is likely to generate inefficiencies. We examine such a market in both theory and practice. The predictions of our model are given strong support by market data.

1. Introduction

Households face many choices that require financial judgment. This judgment relies on numeracy skills and on previous experience with financial products. When financial judgment falls short of the mark, households may make sub-optimal decisions. In some cases, failing to optimize entails a negligible cost; in other cases, this cost is large. Depending on the market in question, agents that make sub-optimal decisions may or may not affect the market equilibrium.

This paper examines certain aspects of a decision faced by many households: making a debt-financed acquisition of a home. We use a survey to identify a consumer bias with regard to debt-financing, and suggest some plausible causes of the bias. Next, we use a simple two-type model to examine how biased consumers may affect prices in equilibrium. Finally, we show that market data support the predictions of our model.

Debt is at the heart of household finance. It allows for Pareto-improving equalization of intertemporal marginal rates of substitution, as households incur debt during less affluent parts of their life-cycle and pay off the debt during more affluent times.

[1] We thank Anna Dreber, Tore Ellingsen, Peter Englund, Christian Ewerhart, Magnus Johannesson, Yannis M. Ioannides, David Laibson, Stephan Meier and seminar participants at the Stockholm School of Economics for helpful comments and suggestions. We thank HSB Riksförbund for sponsoring our survey, Mäklarsamfundet and Värderingsdata for supplying the market data, and Boreda AB for supplying co-op balance sheet data. Johan Almenberg thanks the Jan Wallander and Tom Hedelius Foundation, and the Torsten and Ragnar Söderberg Foundations, for financial support.

The core arguments of the associated optimization problem are the cost of capital and the individual's expected income path. For many households, their home represents the bulk of their assets, and their mortgage contract is the most important financial contract they ever enter into (Campbell and Cocco, 2003). Sweden, which is the focus of our study, is no exception: real estate amounts to over 70% of household assets (Campbell, 2006) and the total amount of mortgages outstanding is equivalent to almost 60% of GDP.[2] Despite this, there has been little research focusing on household mortgage choice. Instead, the existing research into mortgages has largely focused on mortgage-backed securities.

The Swedish market for owner-occupied apartments offers an opportunity to study how households perceive capital structures. A key feature of this market is that all apartments are organized as co-operatives (hereafter: "co-ops").[3] Condominiums are not permitted.[4] The co-ops can, and typically do, take on debt. The cost of servicing the debt is passed on to the co-op members as a part of a monthly or quarterly fee that also includes service and maintenance costs. Since condominiums are prohibited, we avoid the endogeneity problem that arises when comparing co-ops and condominiums in cities or countries that allow both.

In this large market, which represents about one fifth of Swedish households, households debt-finance the purchase of a home in two ways. They can take out a large personal loan and buy a share in a co-op with low leverage, or take out a smaller personal loan and buy a share in a co-op with high leverage. The two sources of debt-financing are substitutes. In theory, under the appropriate conditions, it should not be possible to add value to an apartment by changing its capital structure – analogous to a familiar result in corporate finance (Modigliani and Miller, 1958).

The presence of two sources of debt-financing poses a challenge to household optimization. A consumer who is choosing between different apartments now also makes a choice between different capital structures. Assessing the value of an apartment conditional on the amount of debt held by the co-op is demanding in terms of financial literacy and numeracy, not least since the relative cost of the two sources of debt has varied over time.

Currently, interest payments on individually held debt are tax deductible, whereas interest payments through the co-op are not. This discrepancy makes debt-financing

[2] Source: Swedish Bankers' Association.

[3] Source: Statistics Sweden.

[4] In a condominium, apartments belong to their respective inhabitants. Co-op apartments, by contrast, are owned by a legal entity – the co-op itself – which in turn is owned by the co-op members. The resident owns an exclusive right to inhabit their co-op apartment. Unlike in some other countries, Swedish co-ops have very weak screening rights.

through the co-op almost 50% more expensive than individually held debt. In a frictionless market, debt would immediately be substituted toward the cheaper source. In practice, many households may be slow to optimize co-op capital structures. Debt held by the individual household is arguably more salient – harder to ignore – than debt held collectively by the co-op. The co-op fee is typically not itemized, so in order to ascertain what share of the fee is used to service co-op debt, a co-op member must make his own calculations based on the co-op's annual report. This is demanding in terms of financial literacy.

The effects of cost salience and limited financial literacy do not play a part in the standard intertemporal optimization problem. In practice, however, these factors may have important effects on economic behavior. If households experience disutility from awareness about their degree of indebtedness, the less salient source of debt would carry a non-pecuniary advantage. In this case the value of an apartment may well be affected by its capital structure. In addition, households that find it difficult to calculate the relative advantages of the two sources of debt-financing may underestimate the costs of a sub-optimal capital structure.

Our survey data indicate that many consumers do not have a clear understanding of the optimization problem at hand. We find a pervasive tendency to compartmentalize the two sources of debt-financing. Most survey participants were well-informed about their individually held debt, but poorly informed about the debt held by their co-op. Strikingly, the vast majority have *never* thought of the two forms of debt as being, to some extent, substitutes. If our sample is reasonably representative, this suggests that a sizable fraction of the market is at best characterized by a boundedly rational understanding of the co-op capital structure. These findings are highly consistent with a scenario in which financial literacy, numeracy, and salience jointly produce a systematic bias in the market, in favour of co-op debt. In our view, the difficulty many households have in understanding the co-op capital structure may go a long way in explaining why such debt is still prevalent, despite being considerably more costly.

We have good reasons to expect the bounded rationality of some fraction of market participants to affect prices in equilibrium. There are certain conditions under which individual mistakes need not distort prices in the long run. In financial markets, for examples, arbitrage offers sophisticated investors the opportunity to profit from mis-pricing by naive investors. Alternatively, competition may drive firms to reveal competitors' exploitation of consumers' mistakes, in order to gain market share. Neither of these factors are likely to apply to the case at hand. The Swedish market for apartments is characterized by rent control and other restrictions on renting out apartments. This amounts to a major limitation on arbitrage opportunities. Converting

owner-occupied apartments into rental apartments would typically entail a significant financial loss. When each household owns a single apartment at a time, transactions costs from moving make it unlikely that arbitrage will correct prices, or that any household could internalize the benefits from "educating" the market.

We present a simple two-type model of this market. Both types − naive and sophisticated agents − solve the same optimization problem, but naive agents systematically underestimate the relative cost of co-op debt. Sophisticated agents, by contrast, correctly perceive this relative cost. As a result, naive investors have a more favorable attitude toward capital structures where more of the debt is held by the co-op and less by the individual household.

The model predicts a muted market response to changes in the relative price of the two sources of debt. If naive agents constitute a sizeable fraction of the market – and our survey suggest this is indeed the case – then an increase in the relative price of co-op debt will have a smaller effect on the relative price of high-debt co-ops than if the entire market consisted of sophisticated agents.

We test this hypothesis against market data by using a natural experiment, in the form of a large sample of apartment sales around the time of an unexpected reform in October, 2006. The reform removed a tax that was paid by co-ops with low debt levels, but not by co-ops with high debt levels. As a consequence, the relative cost of servicing co-op debt increased dramatically, making co-op debt almost 50% more costly than individually held debt. In an efficient market, this change should have had a negative effect on the relative price of apartments in high-debt co-ops compared to apartments in low-debt co-ops. Our analysis of sales data shows that while the relative price change for high-debt/low-debt apartments has the right sign, the effect was small and far from statistically significant. Given the large sample size, we conclude that the data provide a strong indication that the market response to this natural experiment was, at best, muted − in line with the prediction of our model.

The rest of the paper is organized as follows. In Section 2, we review some of the related literature. In Section 3, we describe the survey and report its results. In Section 4, we present our model of a market where sophisticated and naive consumers interact in the market for owner-occupied apartments. In Section 5, we test the prediction from the model, using a large sample of apartment sales. Section 6 concludes.

2. Related literature

Our paper examines how the lower salience of co-op debt, together with limited financial literacy, may generate a consumer bias leading to widespread mispricing in the housing market. This line of reasoning fuses several strands of research, primarily

within public economics, behavioral economics, and household finance. In this section, we outline some of the most relevant literature on salience and financial literacy. We briefly discuss under what general circumstances consumer bias may be expected to affect markets in equilibrium. We also review some of recent behavioral studies of the housing market.

2.1. Salience. In economics, the term "salience" is used to emphasize that all costs are not equally transparent. A cost is salient if it is relatively visible, transparent, and hard to ignore. In our model, naive agents are biased toward co-op debt, which is less salient for at least two reasons. First, the responsibility for dealing with the debt is delegated to the co-op board. Second, the debt payments are included in a monthly service charge from the co-op, which is not itemized. Our survey results, presented in section 3, show that individuals are considerably more aware of their personal debt than of the debt held by their co-ops, consistent with co-op debt being less salient.

Salience has been addressed extensively in public economics. Buchanan (1967) conjectures that the government sector may grow in part through finding new, less salient ways of raising revenue. A prominent example is income tax withholding, which makes the taxation of labor less visible to employees, who do not have to make the payments themselves. Another example is social security taxes, which make taxation less lucid by conflating insurance and redistribution. When taxpayers underestimate the cost of government services, demand for government services increases: there is "fiscal illusion".[5]

Empirical work in public economics has shown that tax salience can have large effects on individual behavior. Finkelstein (2007) studies the effects of introducing electronic road toll collection, which is a less salient levy than cash payment. She finds that toll rates increase by about 20-40%, and the short-run elasticity of driving (with regard to toll costs) decreases in absolute value. Chetty et al (2007) find that demand elasticity for alcohol is greater with regard to changes in the excise tax, which is included in the displayed price, than to changes in the sales tax, which is not. In a parallel field experiment, a tax-inclusive price was posted below the ordinary price tag for selected goods in a grocery store. Scanner data revealed an 8% reduction in sales for products with tax-inclusive prices, relative to control products and control stores. And yet a survey conducted at the same time shows that the consumers were well-informed about sales taxes, suggesting that their results are driven by the salience of the tax, and not by ignorance of the tax.

[5] Buchanan attributes this term to Puviani (1903).

2.2. The pain of payment. While the public economics literature on salience has focused largely on taxation, this concept can be applied to any type of cost, including the pre-tax price. A parallel literature in consumer finance looks at how the salience of the payment itself affects consumption. In a series of field experiments, Soman (2006) finds that consumption is increased when payments become less transparent. His interpretation of this finding is that consumers experience a "pain of paying" and that the pain is smaller not just for lesser amounts but also for less apparent payments.[6] Soman uses consumer interviews to rank payment methods by transparency. Cash is considered the most salient, since the money is measured out physically and the exact amount is articulated. Check payments are slightly less salient, but the amount still has to be spelled out in writing. Card payments are deemed to be considerably less salient, and automated payments the least salient.[7]

Empirical evidence from consumer finance research indicates that the salience of payments has real consequences for consumption patterns. Soman (1999) finds that credit card users are more likely to underestimate, or forget entirely, the amount spent on a recent purchase. Prelec and Simester (2001) find that subjects paying by credit card bid more for a prize. The reported effect is large – around 100% – and does not appear to be driven by liquidity constraints.[8] The participants were MBA students at a prestigious business school, making it hard to attribute this result to low levels of financial literacy.

The research on salience in consumer finance points to a deeper behavioral dimension of the consumption decision, namely the voluntary disassociation of payment and consumption. Prelec and Loewenstein (1998) suggest that consumers may devise strategies for optimizing this "decoupling" of costs and benefits. Less salient payment methods facilitate this separation, thus "anaesthetizing" the consumer from the painful payment. This conjecture is supported by Knutson et al (2007), who use functional magnetic resonance imaging (fMRI) to show that the purchasing decision activates two separate regions in the brain. The nucleus accumbens is activated by the anticipated gain (the enjoyment of the good), whereas the insula is activated, and the mesial prefrontal cortex is deactivated, by excessive prices. The latter two regions are associated with anticipated loss, suggesting that consumers experience a real "pain of paying".

---

[6] Soman attributes the term "pain of paying" to Zellermayer (1996).

[7] "Cash is the most transparent form of payment – when one pays by cash, one sees exactly what they are paying. (...) At the opposite extreme, a completely opaque (non-transparent) form of payment might be a payroll deduction that one is not even aware of." (Soman, 2006, p. 175.)

[8] In two of the experiments, the prize was a pair of tickets to a sporting event; in one experiment, a banner. The average bid increased by 113% and 76% respectively, for the tickets, and 59% for the banner, when subjects were told they would pay by credit card. (The variation is between-subject.)

Some financial arrangements may lend themselves more easily to the decoupling of costs and benefits. When consumers choose between co-op apartments, they are also choosing between different capital structures. Prelec and Loewenstein (1998) suggest that changes to the composition of an individual's balance sheet may have hedonic effects even when the economic effects are negligible. It is possible, and in our view plausible, that some consumers add value by choosing a particular capital structure even when the economic benefits are small or negative.

The decoupling of costs and benefits is a particular case of a more general phenomenon: the "mental accounting", whereby individuals compartmentalize elements of their consumption and the associated expenditures into mental accounts, following personal rules or heuristics (Thaler, 1985). This compartmentalization puts limits on the fungibility of money, and gives rise to apparent inefficiencies, as "individuals fail to undertake some internal arbitrage opportunities that in principle could increase utility" (Thaler, 1985, p.212). These inefficiencies stem from the treatment of different elements of consumption and expenditure as belonging to different realms, and not as arguments of a single all-encompassing optimization problem. We argue, on the basis of our survey results, that this is an apt description of consumer behavior in the Swedish market for owner-occupied apartments.

Mental accounting may serve multiple purposes: on the one hand, it can be a tool for self-control that helps boundedly rational individuals make reasonably good judgments (Thaler and Sunstein, 2008); on the other hand, it allows individuals to picture financial matters in a more palatable but potentially inefficient way. There is much room for conflict between the two goals, as an accurate representation is not necessarily the one that makes the individual the most happy in the short run (Prelec and Loewenstein, 1998). In effect, there is a trade-off between economic efficiency and "hedonic efficiency".

The mental accounting model in Prelec and Loewenstein (1998) predicts that individuals would choose to pre-pay for consumption, because debt is particularly unpleasant – what the authors term "debt aversion".[9] But individuals do not always have this option. For housing, the savings required for pre-payment may preclude purchasing a home for a large part of an adult's life. An alternative strategy is to mentally separate a debt from the good for which it was incurred. Kamleitner and Kirchler (2006) study how loan users mentally integrate or separate debt-financed consumption and payment of the debt. They find a one-way connection from the loan to the good: the loan is

[9] "In the traditional economic analysis of consumer choice, consumers are assumed to finance expenditures so as to minimize the present value of payments. (...) The psychological reality of payment decisions is more complicated... The first and perhaps most obvious complication is that debt is unpleasant." (Prelec and Loewenstein, 1998, p.5)

associated with the good, making it more palatable, but the good is not associated with the loan, and hence it is not tainted by the associated debt. The authors describe this as "hedonically efficient". Given the prevalence of such behavior in credit markets, it seems highly plausible that prospective apartment buyers would be attracted to a balance sheet mechanism that allows them to make part of their mortgage less salient.

2.3. Financial literacy. If co-op debt is less salient, and if consumers get disutility from salient debt, then they raise their utility by choosing co-op debt over more salient, individually held debt. If co-op debt is more costly, however, consumers should weigh the psychological benefits of less salient debt against the economic costs. In practice, this entails having a clear understanding of the co-op capital structure and of the current tax breaks. At the most practical level, it also involves some calculations. We have reasons to believe that a non-negligible fraction of consumers will struggle with some, or all, elements of this process. Taken together, the psychological benefits of less salient debt and the inability to assess the costs of a sub-optimal capital structure may generate a consumer bias in favour of co-op debt.

Research on financial literacy sheds more light on the limited ability of many consumers to critically evaluate financial problems or financial products. With regard to consumer finance, two plausible causes of mistakes are (1) a lack of familiarity with certain financial products, and (2) numeracy skills that are inadequate for some financial calculations. A rapidly growing body of research within household finance in general and financial literacy in particular offers insights into both these factors. The findings in this research area make it increasingly difficult to assume that households do not make mistakes, even in relation to relatively simple financial products.

There is ample evidence that many consumers have low levels of numeracy and a poor understanding of financial products. Banks and Oldfield (2006) show that in a large sample of UK citizens close to retirement, a considerable fraction are unable to perform even the most basic financial calculations.[10] In their sample, numeracy is positively correlated with education and negatively correlated with age. Lusardi and Mitchell (2007a) use a similar US sample to evaluate the determinants of wealth close to retirement. Two thirds of the sampled individuals have *never* tried to figure out how much they need to save for their old age. In a broader sample of US consumers, Hilgerth, Hogarth and Beverly (2003) find that less than half of the respondents report using a spending plan or budget. Recent research by the Office for National Statistics indicates that half of the adult population in the UK does not have an understanding

[10] The data in question is the 2002 wave of the English Longitudinal Study of Ageing (ELSA). The numeracy question included in our survey, and reported in section 3.4., is based on one of their questions.

of basic mathematical concepts beyond addition and subtraction (cf. Miles, 2003). Lusardi and Mitchell (2007b) summarize the results from a number of other surveys, and conclude that poor financial literacy is a widespread problem, in the US and elsewhere: "Many households are unfamiliar with even the most basic concepts needed to make sensible saving and investment decisions" (Lusardi and Mitchell, 2007b, p.1). They find large differences between demographic groups, pointing to a link between financial literacy and inequality.[11] Financial literacy has also been found to decline strongly with old age (Laibson et al, 2008; Lusardi and Tufano, 2008).

A striking illustration of how widespread low numeracy might be is the so called "lottery question" used by Banks and Oldfield (2006) and Lusardi and Mitchell (2007a). The question reads as follows: "If 5 people all have the winning number in the lottery and the prize is 2 million dollars, how much will each of them get?" In both samples, only just over half of the respondents were able to answer the question correctly.[12] We use a simplified version of this question in our own survey. As reported in Section 3.4., our results are broadly in line with theirs.

Consumers with a poor understanding of financial products can protect themselves by opting out of various markets. Mounting evidence is suggesting, however, that non-participation is no panacea. For a start, households may not realize that they are ill-equipped to make a particular financial decision: Lusardi and Mitchell (2007b) find strong signs of over-confidence, with households often overestimating their understanding of financial matters. The lack of sophistication of many consumers who choose to participate in financial markets, and the importance of cost salience, is illustrated by a number of recent experiments involving mutual funds. In one experiment, 50 current mutual fund investors were offered a choice of 36 different combinations of load and management fee (Wilcox (2003). Investors typically choose between funds with a highly salient front-end load and no-load funds with higher, but arguably less salient, fees. Wilcox finds that investor choices systematically overweight the more salient front-end loads, in direction contradiction to normative economic theory.[13] In a similar experiment, Choi et al (2006) let subjects allocate an endowment across four S&P 500 index funds with no active management but different fee structures. Subjects were rewarded

[11] The problems of low numeracy skills and low leves of financial literacy is particularly severe for Hispanics, blacks, women, and those with low levels of education.

[12] Banks and Oldfield (2006): 50.8%; Lusardi and Mitchell (2007a): 55.9%.

[13] Specifically: if an investor aims to hold the fund for X years, then a 1 basis point increase in the annual fee is approximately tantamount to an X basis points increase in the load. Wilcox (2003) finds that this rate of substitution for annual fee versus up-front fee is about 2:1, despite a reported average holding time of 17 years. Not one out of 50 participants in the study reported a rate of substitution greater than 3.

based on the subsequent performance of their portfolio.[14] Participants were Harvard undergraduates and Wharton MBA students. One would expect them be financially literate and have high numeracy skills. Despite this, the vast majority inappropriately attached a lot of importance to returns since inception, which for index funds depends only on the inception date. Meanwhile, they neglected the negative impact of high fees: 95% of the participants failed to minimize fees. This is consistent with Alexander et al (1998), who find that in a large survey sample of real-world mutual fund investors, 81% could not give an estimate of the expenses of their mutual funds, not even for the largest mutual fund in their holdings.[15] Only 16% thought that higher fees were negatively correlated with lower net returns, whereas 20% thought the correlation was positive. (At the time, the correlation between fees and net returns was negative. See Alexander et al, 1998, p.310). 40% of those surveyed said they never used the mutual fund prospectus.

Low levels of numeracy and financial literacy can have serious consequences for individual financial well-being. Banks and Oldfield (2006) find a strong correlation between numeracy and pension savings (level and form). They also find a strong positive correlation between numeracy and wealth, and between numeracy and owning shares or owning a private pension. The correlations are robust to a number of controls, including other measures of cognitive ability. The most numerate are also considerably more likely to be well-informed about their pension arrangements. Lusardi and Mitchell (2007a) report a strong correlation between planning for retirement and net wealth. Planning, in turn, is largely carried out by the more financially literate. Lusardi and Tufano (2008) find systematic evidence that individuals with lower levels of financial literacy transact at higher costs. Stango and Zinman (2008) find that individuals who are more prone to payment/interest bias tend to borrow more and save less. In a South African field experiment, Bertrand et al (2005) show that financially irrelevant information can be an important determinant of consumer borrowing. A short-term loan offer differed in the interest rate offered and in terms of other, non-economic framing (what the authors term "psychological manipulation"). The non-economic manipulations had a sizeable average effect on the take-up rate, equivalent to half a percentage point change in the interest rate – implying that consumers that are affected by irrelevant information may be paying a high price for this.

[14] The investments were intermediated by the experimenters and the rewards only based on returns; thus the choice of fund was unbundled from the services provided by the fund.

[15] A majority (57%) also report that they did not know any of the expenses of their largest fund at the time of purchase either.

2.4. Do irrational agents affect equilibrium prices? While individual mistakes clearly affect individual outcomes, it is less evident that irrational actors can affect the market equilibrium. Some market conditions under which this may occur are described by Akerlof and Yellen (1985), who characterize equilibria where the presence of non-maximizing behavior, alongside standard maximization, results in a first-order change in equilibrium values, but only second-order small losses for the non-maximizers. In such cases, they argue, mistakes may go uncorrected. Russell and Thaler (1985) offer a broader description of conditions under which mistakes may prevail in equilibrium: costless arbitrage may not be possible, and evolutionary forces (the selection and propagation of advantageous strategies) may be weaker for households than for firms: "quasi rationality is rarely fatal" (Russel and Thaler, 1985, p.1074). In their view, the classical response, that competition will make irrationality irrelevant for equilibrium, only holds in "very special cases, probably rarely observed in the real world." (Russell and Thaler, 1985, p.1071). We note that rent control prevents large scale arbitrage in the Swedish market for co-op apartments. Moreover, mispricing in this market is unlikely to drive individual households out of the market altogether, i.e., their is little selective pressure on household behavior.

In some cases, market forces will correct for mistakes and reach the rational equilibrium. This might not necessarily be the case, however. For example, if a firm is exploiting a widespread psychological bias, a competitor may inform the public about the bias and hence reveal the exploitation. Under certain market conditions, however, it is not profitable for the competitor to educate the clients of other firms. In this case, de-biasing will not occur – there is a "curse of de-biasing" (Gabaix and Laibson, 2006; see also Russel and Thaler, 1985). Because of the size of the Swedish market for owner-occupied apartments, and the limits on arbitrage, we find it unlikely that any agent would be able to internalize the benefits of educating naive agents in this manner to the extent where it would offset the costs of doing so.

There is mounting empirical evidence of systematic sub-optimization in financial markets. Ausubel (1991) reports a large and persistent discrepancy between credit card rates and the cost of funds. Despite excellent potential to be a competitive market, and thus to function as a continuous spot market for credit, credit card lending rates were highly sticky throughout the 1980s despite large fluctuations in the costs of lendable funds. Search and switch costs appear insufficient to explain the phenomenon: the average consumer in the sample has a $250 surplus to be extracted by switching to a competitive rate. Ausubel also points out that credit card marketers have reported a much greater sensitivity to the more salient annual fee than to a less salient hike in the interest rate.

Stango and Zinman (2008) document a widespread and systematic misunderstanding of the interest rate associated with a given principle and loan repayment schedule – what they term "payment/interest bias". They attribute this to the difficulties many individuals have with exponentiation. The authors emphasize that the use of fuzzy math – heuristics for dealing with complex mathematical problems – is common, and that this just happens to be an instance where the fuzzy math is systematically biased in one direction – toward saving less and borrowing more. In Stango and Zinman (2007) use variation in enforcement of the Truth in Lending Act (TILA) across time and between lenders to show that lenders exploit payment/interest bias by shrouding their interest rates.

Wilson and Waddams Price (2006) find that a large share of consumers do not switch to the cheapest provider in the UK energy market. Aggravating matters further, many of those that switch do so incorrectly, incurring higher costs as a result. This is not easily explained by search costs.

2.5. Behavioral factors in the housing market. It would not be possible to address behavioral research on the housing market without mentioning the work of Robert Shiller (for an overview, see Shiller, 2005). For decades, Schiller has been studying individual attitudes toward the housing market with the assumption that these attitudes need not be rational in a narrow, economic sense. In particular, individuals are prone to both wishful thinking, as in the form of extrapolation of positive price trends, and errors. Shiller stresses that housing prices, in aggregate, need not be efficient.

Some recent empirical work in real estate finance lends support to the hypothesis that irrational actors may affect equilibrium prices in the housing market. This work tests for particular biases that have been documented in economics and psychology, such as loss aversion (Genesove and Mayer, 2001), and money illusion (Brunnermeier and Julliard, 2005).[16]

Both of these studies find that psychological biases have significant explanatory power when it comes to understanding housing market dynamics. Psychological biases may even have more explanatory power than traditional economic variables: Genesove and Mayer note that a large part of the effect of liquidity constraints reported in a previous study disappears when controlling for loss aversion.

A market in which psychological variables are important determinants of behavior is unlikely to be perfectly rational. Brunnermeier and Julliard (2005) explicitly link their analysis to housing market bubbles, in that nominal interest rates closely follow

[16] For experimental/survey evidence and a general discussion of loss aversion and money illusion, see, respectively, Kahneman and Tversky (1991) and Shafir, Diamond, and Tversky (1997).

the trends in mispricing, whereas the real interest rate does not. Moreover, the loss aversion effect reported in Genesove and Mayer (2001) is larger for owner-occupants than for investors. This suggests that the resulting inefficiencies could be even larger in Sweden, where the market for owner-occupied apartments is completely dominated by owner-occupants.

Determining the right price for a house or an apartment is a complicated task, as is the selection of an optimal mortgage. Given this, it is not surprising that in addition to indications of systematic biases, there is also mounting evidence of ignorance and outright mistakes in the housing market. Miles (2003) reports a large share of the UK mortgage market is characterized by bounded rationality and myopia, including a survey that found 10% of respondents did not even know if their mortgage was variable or fixed rate (see Miles, 2003, p.45). Consumers do not have a good understanding of the products, and focus very heavily on the current, moving rate. Risk profiles of mortgage products play a secondary role. In the US, households are slow to refinance their debt in the face of declining interest rates and particularly slow to do so if they have low levels of financial literacy (Campbell, 2006). The costs are large: about 1/5 of households pay more than 2% above the current interest rate. These costs are too high to be easily explained in terms of transaction costs. Campbell (2006) also reports indications that many households choose their mortgage product (fixed rate or adjustable rate) on non-economic grounds.

If individuals struggle with choosing the right mortgage when there is only one source of debt, then we find it highly plausible that mistakes will flourish in a market where there are two sources of debt and co-ops differ with regard to their capital structure.

3. A survey of co-op residents

As a prelude to our analysis of the market for co-op apartments, we conducted a survey of co-op residents that allowed us to better understand how individuals view the apartment financing problem. Participants were asked about their mortgages and about the debt of their co-ops. We also asked them if they had considered the possibility of substituting individually held debt for the co-op debt. They were also asked some questions testing how well they understood the regulations regarding co-ops. In particular, we checked whether they were aware of the tax shield difference between the two sources of financing. We asked some general background questions, including how long they had lived, and were planning to live, in their apartment. We also asked a question intended to check their numeracy skills. The following section focuses on our key questions and the associated findings. We would like to emphasize that the purpose

of the survey was to look for indications of mental accounting, or some other bias, that would lead our sample of co-op residents to have a less-than-fully-rational view of the co-op capital structure. It is beyond the scope of this to identify the determinants of such behavior.

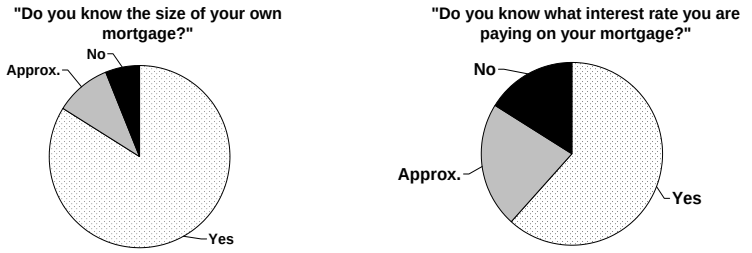
The survey was conducted in February 2008, at the main train station in Stockholm. Participation in the survey was conditional on owning, and being resident in, a co-op apartment. 100 individuals took part in the survey, which lasted approximately 3 minutes. Participants were rewarded with a lottery ticket worth approximately USD4. The mean age of the participants was 45 years, with a minimum of 17 and a mamixum of 77. The sex ratio of participants was exactly 1:1. All but one of the participants had graduated from high school, and about two thirds of the sample had attended university. Approximately one third of the participants had current or previous experience of being on the board of a co-op.

Below we present our four main findings from the survey. (1) Respondents had a high awareness of the size of their mortgage and the interest rate they pay on it. (2) By contrast, respondents had a low awareness of the size of co-op debt, and the interest rate paid by the co-op. (3) Respondents were highly aware of the tax shield differential between individually held debt and co-op debt. (4) Most respondents had never thought about the possibility of substituting individually held debt for co-op debt.

3.1. Debt awareness. We asked survey participants if they knew the size of their own mortgage, and the interest rate they were paying. As shown in Figure 1, the great majority knew the exact size of their mortgage and the exact interest rate they were paying. Of those that did not know the exact numbers, about half knew them approximately. Only 5% of the partcipants in the survey did not even approximately know the size of their mortgage, and only 13% did not even approximately know the interest rate on their mortgage.

Figure 1

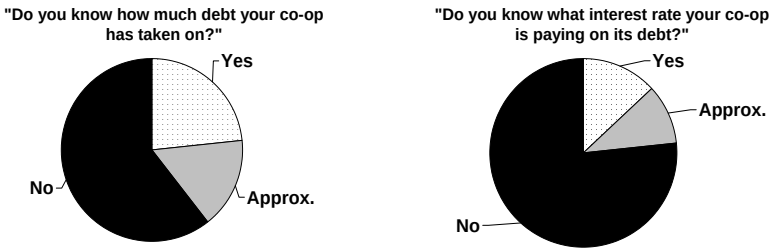
Awareness of own mortgage size and interest rate



By contrast, only a minority of the participants in the survey knew how much debt their co-op had taken on, or what interest rate the co-op was paying. 60% of respondents did not even approximately know how much debt their co-op had, and 76% did not even approximately know the interest rate the co-op was paying.

Figure 2

Awareness of co-op debt size and interest rate



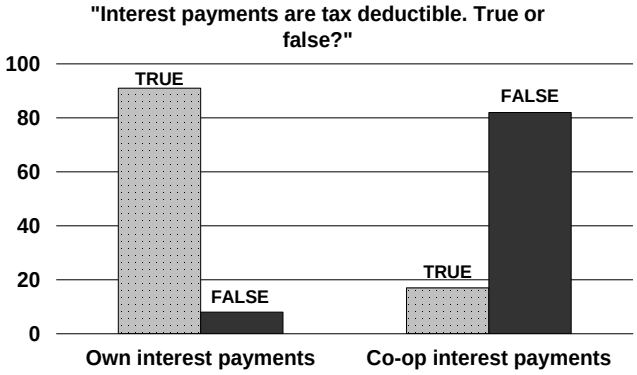
These findings lead us to conclude that debt awareness is considerable higher for individually held debt than for co-op debt, despite the fact that both sources of debt-financing are paid for by the co-op resident. This pattern is entirely consistent with our conjecture that a large fraction of the population may be compartmentalizing co-op debt in a less than fully rational manner. Such mental accounting would imply that for many consumers, co-op debt does not enter into calculations of the value of an apartment in an appropriate way.

3.2. Tax shield awareness. At the time of writing, interest payments on household mortgages are tax deductible. For interest payments up to a specified threshold, 30% of the payment may be deducted from the individual's income tax. For interest payments in excess of this amount, the deduction is 21%.[17] By contrast, interest payments on debt held by a co-op are not tax deductible. The Swedish mortgage market is highly competitive, so the difference between the interest rates faced by co-ops and individuals is negligible. Assuming no difference in gross interest rates, the net cost of capital raised through the co-op is r, whereas the net cost of capital raised through an individually held mortgage is $0.7r$. If individuals are well-informed about the difference with regard to the tax shield, they should seek to reduce co-op debt in favour of individual debt. Debt substitution of this kind has not been occurring to any great extent. In the survey, we investigated the possibility that individuals simply were unaware of the tax shield differential. This does not seem to be the case.

[17] At the time of writing, the threshold is SEK 100,000, equivalent to about USD 12,500.

Figure 3

Awareness of tax shields

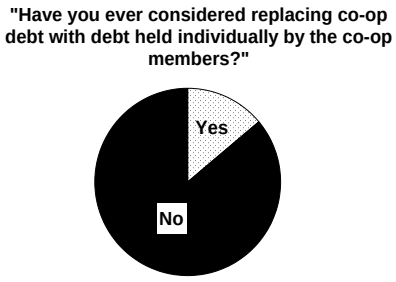


As shown in Figure 3, a great majority of respondents were aware of the fact that interest payments on their own mortgages are tax deductible. The great majority were also aware of the fact that interest payments made through the co-op, as a part of the monthly fee, are *not* tax deductible. 91% correctly identified the first statement as true, and 82% correctly identified the second statement as false.[18] We conclude on the basis of this that there is a high awareness of the tax shield differential between the two sources of financing.

3.3. Awareness of debt substitution. To a large extent, debt held by the co-op and debt held individually by its members are substitutes.[19] In our survey, we wanted to investigate to what extent co-op residents were aware of this.

Figure 4

Awareness of debt substitution



[18] The option "I don't know" was not offered. Uncertain individuals were asked to indicate which response they thought was more likely to be correct.

[19] They are not perfect substitutes, for several reasons. (1) Some households may face credit constraints that prevent them from borrowing directly from the bank. (2) Co-op debt has limited liability in the case of default. The nature of the Swedish mortgage market makes (1) an unlikely explanation for a widespread bias toward co-op debt. Moreover, for loan-to-value ratios of 80% or less (2) is not a convincing explanation. Large price increases in recent years imply that LTV ratios are typically much lower.

As shown in Figure 4, the awareness of the possibility of substituting co-op debt for individually held debt was very low. 86% of respondents had *never* considered the possibility of doing this. Individuals with experience of being on the board of a co-op were more than twice as likely to have thought about this (22% versus 9%). The difference is marginally significant (Pearson chi-squared test; p-value 0.076).

3.4. Numeracy. At the end of the survey, we included the following question: "How much is 2 million divided by 5?". This question is a slightly simplified version of the so called "lottery question" used by Banks and Oldfield (2006) and Lusardi and Mitchell (2007) and discussed in section 2.4. The question is intended to test the respondents' numeracy skills. Respondents were given as much time as they needed to answer the question. If they came up with the wrong answer, but then corrected themselves, we still counted this as a correct answer. Even under these conditions only 67% of respondents were able to answer the question correctly. This figure will probably strike some readers as surprisingly low, but it is actually slightly *higher* than the fraction of correct answers found in the two previous studies that use the lottery question (50.8% and 55.9% respectively). The difference may be explained in part by the slightly simpler phrasing in our survey, or by the fact that participation in our survey was conditional on owning an apartment, which may be positively correlated with numeracy skills.

We wish to emphasize the poignancy of this finding: as we have mentioned, our sample consisted of adults owning, and residing in, co-op apartments. One third of the sample were unable to divide 2 million by 5 without using an aid. In our view, this suggests that many individuals who face quite complicated financial decisions may not have adequate skills for making anything like the optimal decision. Taken together with our other survey findings reported above, this makes it highly plausible that a sizeable fraction of co-op residents will be making sub-optimal financial judgments about co-op capital structures, which is a considerably more complicated problem.[20]

As a word of caution, we note that the survey does not prove that seriously flawed decisions are being made. Nonetheless, we have shown that when given the opportunity to reveal their full understanding of the problem at hand, a reasonably representative sample of apartment-owners failed to do anything of the sort. Rather, the survey results are highly consistent with our conjecture that a considerable share of market participants do not have an economically sound understanding of co-op capital structures.

[20] Our measure of numeracy is positively correlated with awareness of the substitutability of debt, but the correlation (0.16) is not very strong.

4. A simple model of the co-op market

4.1. Modeling Consumer Bias. Mental accounting and low levels of financial literacy can give rise to economic behavior that departs from the predictions of standard economic models. This raises the question of whether a formal model should aim to take such behavioral considerations into account – and if so, how? One crucial distinction is whether such behavior is deemed to be irrational. In economics, behavior is typically described as rational if it confirms to the predictions of normative theory, i.e., economic theory that states what the individual ought to do. But economic modelling itself does not require such rationality. According to one influential definition, a positive theory of economic decision-making is one which generates good predictions regarding economic behavior (Friedman, 1953).

When the departures from normative theory are systematic, the behavior is "quasi-rational" in the sense of being predictably irrational (Thaler, 1980; Russell and Thaler, 1985; Ariely 2008). A widespread psychological bias is exactly the kind of systematic deviation from normative prediction that lends itself to economic modelling. Such models should be thought of as extensions of, and complements to, standard normative models.

We have reason to believe that there is a systematic bias in the market for owner-occupied apartments, in favour of co-op debt. This debt has the advantage of being less salient, which may give rise to psychological benefits. Calculating the cost of choosing co-op debt over individually held debt is demanding in terms of financial literacy. This makes it unlikely that all households fully understand the costs of co-op debt. Together these factors are likely to result in a biased perception of the relative cost of the two sources of debt financing.

One approach for modelling quasi-rational behavior is to treat it as a mistake that occurs when the agent converts raw information into a budget set (Russell and Thaler, 1985). This allows for a distinction between individual differences in (1) preferences, (2) information, and (3) the mapping from the real world to the mental representation of a budget constraint. Having arrived at a not-quite-accurate budget set, the agent optimizes in the same way that a fully rational agent would.

The model outlines in this section follows a similar approach: we assume that some buyers – the "naive" agents – have a biased perception of the relative cost of the two sources of debt financing. In the model, the bias takes the form of an additional psychological cost associated with individually held debt. Aside from this, both naive and sophisticated buyers solve the same optimization problem.

4.2. A model with two types. There is a continuum of agents on $[0, 1]$ that each decide on purchasing an apartment. For simplicity, we assume that all agents finance their purchases through debt. There are two sources of debt-financing, individually held debt and co-op debt, and we allow for convex combinations of the two. Throughout the following analysis, we will refer to individually held debt as *equity* and to debt held through the co-op as *debt*.

Let the gross interest rate r be the same for both equity and debt. Suppose that the cost of capital is tax deductible for equity but not for debt, and that all agents understand this difference when considering the optimal capital structure. Letting τ denote the tax rate, the net cost of a unit of capital raised as equity is $(1 - \tau)r$.

In essence, all agents solve the same maximization problem, but some agents use an incorrect mapping from the information set to the budget set. They get disutility from the higher salience of equity, and fail to perceive the economic costs of the resulting sub-optimal capital structure.

We model this bias as a psychological cost c_j that is proportional to the interest paid on equity. Fraction α of all agents are *sophisticated* (type $j = 1$) and do not perceive a psychological cost. Fraction $1 - \alpha$ are *naive* (type $j = 2$) and do perceive a psychological cost.[21] The net benefit b_j associated with equity is the tax shield less the psychological cost, where the cost is modeled as a convex function of the amont of equity E:

$$b_j E \equiv r\tau E - r(1 - \tau)c_j E^2 \tag{4.1}$$

where $c_1 = 0$ and $c_2 > 0$.

The initial purchase is financed through a combination of equity and debt, i.e., the agents chose capital structures for their apartments. Normalizing the sum of debt and equity to 1, we can denote the capital structure associated with a given equity level as $\{E, 1 - E\}$, where E represents the fraction of equity.

Because the net cost of equity differs from the net cost of debt, the price of an apartment depends on its capital structure. The price is given by the function $P(E)$. Agents are not financially constrained.[22]

[21] Alternatively, this could be modeled as a psychological *benefit* from the less salient co-op debt. Since we are only concerned with the relative cost of the two debt sources, such an approach is equivalent to ours.

[22] We will abstract from the following aspects: (1) Co-op screening. The co-op has veto rights over new members, but in Sweden these rights are very weak. (2) Pros and cons of the co-op versus other forms of ownership. In Sweden, owner-occupied apartments are without exception organized as co-ops. The condominium ownership structure is not permitted. Thus, we take the co-op ownership structure as exogenously given. (3) Default. Rising prices over the last decade have resulted in lower LTV-ratios for co-ops. When the leverage is moderate, default is highly unlikely. For this reason we will abstract from the difference between the two forms of financing in the case of default.

Once the initial capital structure is chosen it cannot be changed. Each type chooses the capital structure that maximizes their benefits throughout their lives, taking into account the possibility that they will sell the apartment in the future to an agent of a different type. Thus the relative supply of apartments with the two different equity structures is determined by the interaction of the two types of agent in the marketplace.

Sophisticated agents choose the capital structure $\{E_h, 1 - E_h\}$, and naive agents choose $\{E_l, 1 - E_l\}$. We show that $E_h > E_l$. The relative supply of apartments with high and low equity structures is determined by α.

In order to choose an apartment, agents make $m = 2$ searches at the beginning of period 2. We assume that when searching an agent views one apartment of either capital structure type. Our results do not hinge on this assumption - all we need is that apartments are sufficiently idiosyncratic that a full separation, whereby either agent type trades only with itself, does not occur. This realistic feautre of the housing market is captured through the (separately additive) idiosyncratic utility v_i, uniformly distributed on $[-V, V]$, that an agent receives when matched with an apartment.[23] For simplicity, we assume that agents do not anticipate this in the first period.

Agents live for two periods. At the beginning of period 1 they make their initial purchase for which they choose a capital structure that maximizes their benefits throughout their lives. After living in the apartment for one period, agents search for new apartments at the beginning of period 2. If they find an apartment that suits them even better they trade. If not, they remain in the same apartment in the second period.

4.3. Determining the price. We solve for the equilibrium by backward induction. First, we see how agents decide on trading their apartments once they have chosen capital structures. Next, we solve for the first period, in which agents choose capital structures anticipating their likelihood of moving to a new apartment and the future resale value of their initial purchase.

4.3.1. ***Period 2: Trading***. The equilibrium price equates supply and demand for apartments of both types of capital structure. The supply of E_h and E_l apartments depends on α and is determined in period 1. Below, we calculate the demand in several steps.

By assumption each agent considers one apartment with high equity and one with low equity. Both types may settle for an apartment with high or low equity, whichever gives them the higher utility given the realization of the random variable v. Suppose for simplicity that the individual utility of monetary payoff M is $u_i(M) = M$. Then an agent of type j will trade their initial purchase for an apartment of equity level E_h

[23] The model is tractable for a general number of searches, but this does not offer further insight.

if

$$v_i^a - P_h + (E_h)b_j > v_0 - P_0 + (E_0)b_j \tag{4.2}$$

and

$$v_i^a - P_h + (E_h)b_j > v_i^{a'} - P_l + (E_l)b_j \tag{4.3}$$

where a' denotes the low equity apartment, and a the high equity apartment. P_h and P_l are the prices of high and low equity apartments, respectively. The parameter b_j is the net benefit of equity as defined above. It includes both the psychological cost and the value of the tax shield. Condition (4.2) states that the agent will move to a new apartment if the sum of the utility from living in the new apartment and the net benefits associated with its capital structure, less the cost P_h of purchasing the new apartment, is greater than what the agent earns by staying in the initial purchase (subscripts 0).

The probability that a sophisticated agent trades a high equity apartment for another high equity apartment in the second period is

$$P_1^h \equiv P\{v_i^a - (P_h) + E_h b_1 > v_i^{a'} - (P_l) + E_l b_1\} \times P\{v_i^a > v^0\} \tag{4.4}$$

where the last term is calculated from (1) for type $j = 1$.

Similarly, the probability that the agent will stay in their current E_h apartment is

$$P_1^0 \equiv P\{v^0 - (P_h) + E_h b_1 > v_i^{a'} - (P_l) + E_l b_1\} \times P\{v_i^a < v^0\} \tag{4.5}$$

The total demand by agents of type $j = 1$ for high equity apartments is $\Pi_1^h = P_1^h + P_1^0$.

The probability that naive agents will move from low equity apartments to high equity apartments can be calculated in the same manner. The equilibrium price equates supply and demand in the market for high equity apartments:

$$\alpha\Pi_1^h + (1-\alpha)\Pi_2^h = \alpha \tag{4.6}$$

where α is the proportion of sophisticated agents, as before. Similarly, the equilibrium price equates supply and demand for the low equity apartments:

$$\alpha\Pi_1^l + (1-\alpha)\Pi_2^l = 1 - \alpha \tag{4.7}$$

where $\Pi_2^l = P_2^l + P_2^0$, analogous to the total demand for high equity apartments.

PROPOSITION 1. *The equilibrium price of high equity apartments is given by*

$$P_h = P_l - 2\alpha V + V + (E_h - E_l)\big(rt - (1-\alpha)c_2(1-r)\tau(E_h + E_l)\big) \tag{4.8}$$

Note that the price is decreasing in the psychological cost c_2, which causes the price of high equity apartments to be lower than the price of low equity apartments.

Proof: See Appendix A.

This is consistent with our survey finding that many agents indeed do not consider substituting equity for debt, even though they understand the tax advantages of equity. The prevalence of such behavior in the market decreases the price of high equity apartments compared to what one would observe given the favorable tax treatment of equity relative to debt.[24]

PROPOSITION 2. *The price response to a change in the tax advantage of equity relative to debt is smaller, in absolute terms, when some agents are naive.*

Proof: In an economy without naive agents, i.e., when $c = 0$ and/or $\alpha = 1$,

$$\frac{\delta(P_h - P_l)}{\delta\tau} = (E_h - E_l)r \tag{4.9}$$

which is larger, in absolute terms, than the response that the model predicts when $c_j > 0$ for at least one type j, and $\alpha < 1$:

$$\frac{\delta(P_h - P_l)}{\delta\tau} = -(E_h - E_l)\big(r - (1-\alpha)c(1-r)(E_h + E_l)\big) > -(E_h - E_l)r \tag{4.10}$$

In other words, if there are naive agents in the market, then we expect equilibrium prices to be less responsive to a change in the relative cost of the two sources of financing. We explore this hypothesis in the first empirical analysis section.

4.3.2. ***Period 1: choosing capital structure.*** When choosing capital structures, agents maximize their tax benefits, less the psychological cost of equity, and the expected gain or loss from trading in the second period. In addition, each type of the agent chooses an optimal capital structure taking into account the optimal choice of the other type.

PROPOSITION 3. *Naive agents choose a lower level of equity, relative to sophisticated agents, when determining their capital structures in period 1.*

Proof Naive agents weigh the tax benefits of equity against the psychological cost:

$$\max_E r\tau E - rE - (1-E)r - cE^2 \tag{4.11}$$

From the first-order condition we get the optimal level of equity for naive agents:

$$E_l = \frac{r\tau}{2c} \tag{4.12}$$

while sophisticated agents prefer not to hold any debt at all, so $E_h = 1, D_h = 0$[25]

[24] Note that the price differential is not necessarily decreasing with α.

[25] One could internalize the optimal choice of equity by imposing a nonlinear cost of raising it. While such an assumption would be innocuous and more realistic, it does not add insight in our model, since our interest lies in the workings of the trade in period 2 for *any* high and low equity apartments chosen by agents.

4.3.3. ***Extension: a dynamic model.***

PROPOSITION 4. *If naive agents dominate the market, an increase in the psychological cost parameter c2 results in an decreased probability of moving. This effect is stronger for naive agents. The probability of staying in the current apartment is given by*

$$P_1^0 = \frac{V + x_1^h}{4V} = \frac{2V\alpha + (1-\alpha)c(1-r)t(E_h^2 - E_l^2)}{4V} \tag{4.13}$$

for type $j = 1$*, and*

$$P_2^0 = 0.5\big(1 - \frac{V + x_2^h}{2V}\big) = \frac{2V(1-\alpha) + \alpha c(1-r)t(E_h^2 - E_l^2)}{4V} \tag{4.14}$$

for type $j = 2$*. Sophisticated agents are more likely to stay in their current apartments because of the naive agents psychological cost if the naive agents dominate (*$\alpha < 0.5$*), because*

$$\frac{\partial P_1^0}{\partial c} \geq \frac{\partial P_2^0}{\partial c} \tag{4.15}$$

Proof See Appendix A.

The intuition for this result is that when a large number of prospective buyers dislike the high equity apartments, sophisticated agents will not receive enough compensation from selling in relation to the expected gains. In a dynamic model, this translates into sophisticated agents staying longer in their apartments when the market is dominated by naive agents. Indeed, if apartments can be traded in every period, the lower probability of trade in each period means that the likelihood that an agent finds an appropriate apartment for trade is lower with any (finite) number of periods. We test the conjecture that sophisticated agents stay longer in their apartments in the second empirical analysis section.

5. Empirical analysis

5.1. Sales data. If all agents in the market are sophisticated, co-op debt should be fully capitalized in apartment prices. Increasing an apartments share of co-op debt should lower the price by the same amount. We conjecture that this is not an accurate description of the Swedish market for co-op apartments. Our survey findings suggest that a boundedly rational view of co-op capital structures may be pervasive in the market. In this section we use a large sample of market data to shed more light on this conjecture.

Due to the particular legal status of Swedish co-ops, there is no centralized, publically available data on their capital structures. Instead, we use the monthly fee, which we can observe in our data set, as a proxy for co-op debt. While heterogeneity with

regard to the service flow provided by co-ops makes this proxy variable less than perfect, we have reason to believe that it is still reasonably good. There are a few firms in Sweden that specialize in analyzing co-op annual reports, including the balance sheets. One of these firms, Boreda AB, allowed us to estimate the correlation between the fee/m^2 and co-op debt/m^2 on their data for 2007. The correlation coefficient for this subsample is about 0.40. When controlling for year of construction (as we do in the following analysis), the correlation rises to about 0.66. On the basis of this, we assume that the correlation between the monthly fee/m^2 and debt/m^2 is reasonably strong also in the national sample of sales data used below.

Since we do not directly observe debt, we cannot measure how debt affects prices by simply regressing price/m^2 on fee/m^2. If the service flow in the co-op is positively correlated with the co-op fee, then the effect of a higher fee on prices is ambiguous. As a result, a regression coefficient on the co-op fee does not lend itself to a straight-forward interpretation.

Instead, we make use of a natural experiment to examine whether co-op debt is capitalized in a rational manner. On 16 October 2006, the government unexpectedly announced that it would abolish a supplementary housing tax levied on co-ops. At the time, all co-ops were required to pay a basic housing tax amounting to 0.5 percent of the assessed value of the property. In addition to the basic housing tax, co-ops paid a 28 percent tax on the imputed rent, calculated as 3 percent of the assessed value, *less co-op interest payments*. The enactment date of the tax change, 1 January 2007, was disclosed at the time of the announcement.

The crucial aspect of the reform is that had different effects for co-ops with high and low leverage. Interest rate payments made by the co-op were deductible against this tax. Hence co-ops with high fractions of debt did not pay the tax, and would not be affected by the reform. By contrast, co-ops with little or no debt would enjoy considerable cost reductions as soon as the reform was implemented at the end of the year.

Forward-looking, rational consumers should have anticipated that co-ops with low leverage would either (1) reduce their monthly charges in the future, or (2) maintain the same monthly charge but increase the flow of services. Thus, in an efficient market we would expect a positive price effect for apartments in low debt co-ops. To separate this effect from general price movements, we focus our attention on changes in the *relative* price of low debt and high debt co-ops around the time of the reform. In an efficient market, this relative price should have changed in favour of co-ops with low levels of debt.

We test this hypothesis using market data on apartment sales from the fourth quarter of 2006, around the time of the reform. The data was provided by the Swedish association of real estate agents, and contains detailed information on more than three thousand apartment sales that took place throughout the country within a $+/-$ 30 day window of the reform. The data distinguishes between the transaction date and the moving date. We use the former to divide the sample into pre-reform and post-reform sales. The data also allows us to control for a number of apartment characteristics, including the monthly fee, apartment size, the number of rooms, what floor, the age of the building, whether the building has an elevator, and the location. Location is measured at the municipal level.

5.2. The econometric model. We fit a hedonic model with the sales price as the dependent variable. We include the co-op fee per square meter as an explanatory variable, as well as a dummy variable for being a post-reform sale, and an interaction term for the co-op fee and the post-reform dummy. The fee, which we can observe, is a proxy for co-op debt, which we do not observe. Co-ops with higher fees are assumed, *ceteris paribus*, to have higher debt levels. We can write our econometric model as

$$y_i = \delta_0 + \beta_1 fee_i + \beta_2 post_i + \beta_3 fee_i \times post_i + \mathbf{X}'\boldsymbol{\delta}_i + \varepsilon_i \tag{5.1}$$

where the dependent variable, y, is the sales price (per m^2), fee is the (annualized, per m^2) co-op fee, *post* is a dummy variable indicating a post-reform sale, and $\mathbf{X}$ is a vector of apartment characteristics.

We are primarily interested in the coefficient β_3. More debt should result in a lower price both before and after the reform, but the negative effect of co-op debt should be larger after the reform, which made such debt more costly relative to equity. The coefficient β_3 captures this difference-in-difference.

5.3. A rational benchmark. In this section, we estimate a benchmark value for β_3 in a market where all agents are rational. For a rational agent, the utility flow from an apartment is not affected by the co-op's capital structure. Hence the flow cost that a rational individual is willing to pay for the apartment should be the same regardless of the capital structure. Letting V, D, and E denote the value of assets, debt and equity, and a, b, and c their respective cost, we can write this condition as

$$\overline{a}\overline{V} = \overline{b}D + \overline{c}E$$

$$\Rightarrow E = \frac{\overline{a}}{\overline{c}}\overline{V} - \frac{\overline{b}}{\overline{c}}D$$

The first derivative of E w.r.t. D gives the marginal rate of substitution between debt and equity such that the agent is indifferent between different capital structures:

$$MRS_{D,E} = \frac{\partial E}{\partial D} = -\frac{\bar{b}}{\bar{c}}$$

Before the reform, $b_0 \approx c_0 = (1-\tau)r \Rightarrow MRS_0 \approx 1$. After the reform, $b_1 > c_1 \Rightarrow MRS_1 > 1$. More precisely, $b_1 = r$ whereas $c_1 = (1-\tau)r$ as before, implying

$$MRS_1 = \frac{r}{(1-\tau)r} = \frac{1}{(1-\tau)}$$

In other words, the post-reform marginal rate of subsitution for a rational individual is such that a marginal increase of one unit of debt reduces equity by $1/(1-\tau)$ units. The fee, however, is a proxy for the interest payment on the debt, rD, not the principal D. We must adjust for this when estimating the coefficient β_3. In the simplest of worlds, a marginal increase of one unit of interest paid implies an additional $1/r$ units of debt. It follows that the marginal rate of substition between interest payments on co-op debt and the level of equity is given by

$$MRS_{rD,E} = \frac{1}{(1-\tau)} \times \frac{1}{r} = \frac{1}{(1-\tau)r}$$

We know that $\tau = 0.3$. Reasonable estimates of r for the second half of 2006 are in the range of 0.05, approximately corresponding to STIBOR + 2%. Table 1 below shows a range of estimates of our rational benchmark, i.e., the effect size if the market responded efficiently to the reform.

Table 1

Estimating a rational benchmark

r	1/(1-τ)	Pre-reform $MRS_0 = 1/r$	Post-reform $MRS_1 = 1/r * 1/(1-\tau)$	Rational benchmark $\beta_3 = MRS_1 - MRS_0$
0.04	1.43	25.00	35.71	10.71
0.05	1.43	20.00	28.57	8.57
0.06	1.43	16.67	23.81	7.14

5.4. Regression results. We fit equation (5.1) to sales data within a window of +/− 30 days of the announcement of the reform. In addition, we fit the model to more narrow windows of 10 and 20 days. A more narrow window reduces the scope for other, background variables, e.g., macroeconomic variables, to bias the results, but also results in a smaller sample. The estimates for the key variables are reported in Table 2.

Table 2

Effect of fee on price per m^2

Variable	Window:	(1) 10 days	(2) 20 days	(3) 30 days
Co-op fee		-21.22	-19.50	-18.37
		(0.000)***	(0.000)***	(0.000)***
Co-op fee*post-reform		2.71	-1.18	-1.84
		(0.383)	(0.583)	(0.298)
Post-reform		-1,506.27	982.86	1,338.56
		(0.420)	(0.449)	(0.213)
Additional controls		Yes	Yes	Yes
Constant		61,472.20	59,583.03	56,795.03
		(0.000)***	(0.000)***	(0.000)***
Observations		1105	2182	3052
R^2		0.84	0.83	0.83

Robust p-values in parentheses. **significant at 5%; ***significant at 1%

For the 10-day window the coefficient on the interaction term, our key variable of interest, has the wrong sign and is not statistically significant. For the 20 and 30 days windows, the coefficient on the interaction term has the appropriate, negative sign, but despite the large sample size and high R-squared, it remains far from statistically significant. This suggests that the market did not react efficiently to the reform, in the sense of prices that incorporate all available public information. This finding is consistent with our survey results, which suggest that a large fraction of the market may not have a good understanding of the co-op capital structure.

Consistent with the prediction of our model, we are unable to reject the null hypothesis that there was no relative price change following the reform. The sales data for the period preceding and immediately following the reform do not indicate that market capitalizes co-op debt in an efficient manner. While it is not possible to prove a negative, the large sample size and good explanatory power of our regressions using the sales data suggest that what efficiency might be in the market is weak at best.

5.5. Robustness check: supply effects. Price is set by demand *and* supply. It is possible that supply changes affected the equilibrium price, perhaps offsetting a change in demand. We check this by comparing the mean and variance of fees observed during the 30 days preceding the reform with those observed during the 30 days immediately following the reform.

As table 3 shows, there was little difference between the two periods with regard to the mean and variance of the annual fee variable. The mean declined slightly and the variance increased marginally. A two-sample t-test is unable to reject the null

hypothesis of equal means (p-value: 0.52). The equality of variance between the two periods is tested using the Levene statistic. We are unable to reject the null hypothesis of equal variance (p-value: 0.99).

Table 3

Testing the equality of means and standard deviations within 30 days before and after the reform

	Pre-reform	Post-reform	Statistic	H_0	p-value
Mean	600.5	597.3	t-test	Equality of means	0.23
Standard deviation	139.7	141.0	Levene	Ratio of standard deviations = 1	0.50

5.6. Survey data. The extended version of our model predicts that for a range of α, sophisticated agents will be less mobile. Our survey data offers an opportunity to test this hypothesis. One of the survey questions was how many years the respondent expected to reside in their current apartment. Responses ranged from less to one year to "for the rest of my life".[26] In order to examine whether duration of residence was correlated with sophistication regarding the capital structure, we needed some measure of sophistication. The survey also included a question about whether the individual had ever thought about the two forms of debt as being, to some extent, substitutes. We use a positive answer to this question as a proxy for being sophisticated.

There is a considerable difference between naives and sophisticates in our sample with regard to how long they expect to stay in their apartments. The former expect, on average, to stay for 7.2 more years, whereas the latter expect to stay more than twice as long, 14.7 years. Our sample is small: only 14 of the 100 respondents replied positively to the question about debt substitution. Moreover, the apparent correlation could be driven by other variables. To shed more light on the statistical significance of this finding and on the role of other variables, we regress the expected residency on a number of controls, including age, sex and education. We also control for how long the repondent had already lived at their apartment. We can write this regression as

$$y_i = \beta_0 + \beta_1 residency_i + \beta_2 age_i + \beta_2 sex + \beta_2 education + \varepsilon_i \tag{5.2}$$

Table 4 presents the results from fitting equation (5.2) to our survey sample. Column (1) shows the estimated correlation from a simple regression without controls, and column (2) shows the regression estimates when we include the controls.

[26] In the latter case, a numerical value was imputed according to the formula A=B-C, where A is the imputed value, B is a constant representing a "normal" expected lifespan, and C is the respondents age at the time of the survey. For the results presented below, B=85. Additional regression with B=80 and B=90 can be found in Appendix X. Our analysis is not particularly sensitive to the selected value of B, although for B=90 the p-value is a little over 0.05.

Table 4

Sophistication is positively correlated with expected length of stay

	(1)	(2)
Sophisticated	7.51	7.01
	(0.010)**	(0.043)**
Years of residency so far		-0.23
		(0.060)*
Age		0.30
		(0.000)***
Male		-0.92
		(0.636)
University		-1.37
		(0.461)
Constant	7.21	-1.94
	(0.000)***	(0.514)
Observations	100	100
R^2	0.09	0.24

Robust *p*-values in parenthesis. * significant at 10%; ** significant at 5%; *** significant at 1%

In the simple regression without controls, the (positive) coefficient on "Sophisticated" is highly statistically significant. When we include the control variables, the explanatory power increases considerably, but the coefficient on "Sophisticated" remains similar in magnitude - implying that sophisticated individuals stay for about 7 years longer, on average - and continues to be statistically significant. While one should be cautious when drawing inference from such a small sample, this finding nevertheless provides support for the model's prediction.

6. Discussion

We have used a survey to identify a consumer bias and shown that it can have economically significant effects for individual welfare as well as equilibrium prices, in both theory and practice. The bias in question is the tendency to compartmentalize two closely related sources of debt as if they were not part of the same financing problem. Looking at the market for owner-occupied apartments, we argue that the lower salience of co-op debt, in combination with low levels of financial literacy, gives rise to a bias in favour of co-op debt. Because many households do not fully comprehend co-op capital structures, they do not fully perceive the costs of failing to optimize the composition of their debt. As a result, the distortionary effects of lower salience are left largely unchecked.

The bias gives rise to sub-optimal capital structures, and may thereby impose substantial costs on individual households. We offer the following numerical example. Following the 2006 reform, co-op debt became almost 50% more costly than individually held debt. In 2007, the average price of a Stockholm apartment was about SEK 2 million. Suppose an apartment trades at this price, and the co-op has a debt/equity ratio of 1:2. Then the co-op debt associated with the apartment amounts to SEK 1 million. At current interest rates, a household with SEK 1 million in co-op debt could reduce their monthy lending costs by about SEK 1,250 by replacing co-op debt with household debt, equivalent to about 6% of the average pre-tax monthly wage.[27]

The prevalence of biased consumers appears to distort market prices in equilibrium. We examine market data before and after an unexpected reform in 2006. If the market was perfectly efficient, we would expect a relative price change in favour of co-ops with low leverage. Despite the large sample size and high explanatory power of our regression, we do not observe this. We observe an effect with the correct sign, but it is small in size and far from statistically significant. The model presented in in Section 4 predicted a muted response to a change in the relative cost of the two forms of debt. Our econometric results are consistent with this prediction.

A bias that leads to sub-optimal capital structures and the mispricing of co-op debt in the housing market may be aggravated by consumers' sluggishness to respond to changes in relative prices. The tendency to adhere disproportionately to the choice that does not require active change has been described as "status quo bias" (Samuelson and Zeckhauser, 1988). When households are uncertain about what constitutes the optimal strategy they may turn to their surroundings for guidance. This can exacerbate the sluggishness: "[U]nsophisticated households tend to use whatever financial contracts are standard in a particular country, possibly because they follow the lead of relatives and neighbors." (Campbell, 2006, p. 34). Herding behavior of this kind would also explain why different types of mortgages are standard in different countries. For example, fixed-rate mortgages constitute the vast majority of mortgages issued in the US, yet hardly exist in the UK (Miles, 2003).

What economic policies might mitigate the bias and its effects on market equilibrium? It is quite possible that regulation could go some way in reducing the scope for costly mistakes in this market. Regulation, however, often comes at the cost of imposing restrictions on all participants in the market. It is important to consider both costs and benefits of different policy options. To give an example, the problem of sub-optimal capital structures could clearly be dealt with by simply banning co-op leverage. We

[27] Source: Statistics Sweden. SEK 18,000 $\approx$ USD 2,200. The example is based on an interest rate of 5%.

believe this to be an unwise policy choice. Short term debt is a convenient way for co-ops to distribute unforeseen expenditures, such as the need to adjust the premises to fit new building laws, over slightly longer time periods. Banning such debt might protect naive consumers, but also imposes an inconvenience cost on all consumers in the market.

Ideally, regulation should make it easier for consumers to make economically sensible decisions while still leaving the final choice to the consumer. This way regulation is kept at a minimally intrusive level. This regulatory approach to household finance has been described as "libertarian paternalism" (Thaler and Sunstein, 2003, 2008) or "asymmetric paternalism" (Camerer et al, 2003). The aim is to target naive consumers without imposing welfare-reducing restrictions on sophisticated consumers.

We suggest two policies that would reduce the scope for mistakes without significant infringements on consumer choice. First, the co-op monthly fee could be itemized so that it is readily apparent what fraction of the fee is used for service and maintenance on the one hand, and interest payments on the other hand. This increases the salience of co-op debt and makes the capital structure more transparent. Second, real estate agents could be encouraged to disclose the co-op debt associated with an apartment in their advertisements. This information can be inferred from a co-op's annual statement. It is unlikely, however, that consumers will read annual statements and make the necessary calculations at the early stages of choosing an apartment. Proving the information in the advertisement itself would serve as a timely reminder to the consumer that the value of assets is the sum of equity and debt, and facilitate quick comparisons between apartments with different capital structures.

Both policies would be inexpensive to implement and would help consumers make more informed decisions. We believe that both the first and second suggestion would be well suited to field experiments, and encourage further research along these lines. There is also a third, far more costly, policy option that has the potential to greatly reduce the cost of failing to understand co-op capital structures: making interest payments on co-op debt tax deductible on par with individually held debt. Aside from political economy considerations, it is hard to see an economic rationale for the current asymmetry. Finally, we note that some existing regulations − notably, rent control and other restrictions on letting co-op apartments − exacerbate the effects of the bias on market equilibrium, by preventing arbitrage by sophisticated consumers. A discussion of the merits of rent control are beyond the scope of this paper, but we recommend that our findings be taken into account in a cost-benefit analysis of the current system.

7. References

Akerlof, G. and Yellen, J. (1985), "Can Small Deviations from Rationality Make Significant Differences to Economic Equilibria?". *American Economic Review* 75(4), 708-720.

Banks, J. and Oldfield, Z. (2006), "Understanding Pensions: Cognitive Function, Numerical Ability and Retirement Savings". Institute for Fiscal Studies Working Paper no. 0605.

Bertrand, M., Karlin, D.,Mullainathan, S., Shafir, E. and Zinman, J. (2005), "What's Psychology Worth? A Field Experiment in the Consumer Credit Market". NBER Working Paper no. 11892.

Brunnermeier, M. K. and Julliard, C. (2008), "Money Illusion and Housing Frenzies". *Review of Financial Studies* 21(1), 135-180.

Buchanan, J. M. (1967), *Public Finance in Democratic Process: Fiscal Institutions and Individual Choice.* University of North Carolina Press.

Camerer, C., Issacharoff, S., Loewenstein, G., O'Donoghue, T. and Rabin, M. (2003), "Regulation for Conservatives: Behavioral Economics and the Case for "Asymmetric Paternalism". *University of Pennsylvania Law Review* 151, 1211-1254.

Campbell, J. Y. (2006), "Household Finance". *Journal of Finance* 61(4), 1553-1604.

Campbell, J. Y. and Cocco, J. F. (2003), "Household Risk Management and Optimal Mortgage Choice". *Quarterly Journal of Economics* 118(4), 1449-1494.

Chetty, R., Looney, A. and Kroft, K. (2008), "Salience and Taxation: Theory and Evidence". Forthcoming in *American Economic Review.*

Choi, J. J., Laibson, D. and Madrian, B. C. (2008), "Why Does the Law of One Price Fail? An Experiment on Index Mutual Funds". NBER Working Paper no. 12261.

Finkelstein, A. (2007), "E-Ztax: Tax Salience and Tax Rates". NBER Working Paper no. 12924.

Gabaix, X. and Laibson, D. (2006), "Shrouded Attributes, Consumer Myopia, and Information Suppression in Competitive Markets". *Quarterly Journal of Economics* 121(2), 505-540.

Genesove, D. and Mayer, C. (2001), "Loss and Aversion and Seller Behavior". *Quarterly Journal of Economics* 116(4), 1233-1260.

Hilgert, M. A., Hogarth, J. A. and Beverly, S. G. (2003), "Household Financial Management: The Connection between Knowledge and Behavior". *Federal Reserve Bulletin*, 309-322.

Tversky, A. and Kahneman, D. (1991), "Loss Aversion in Riskless Choice: A Reference-Dependent Model". *Quarterly Journal of Economics* 106(4), 1039-1061.

Kamleitner, B. and Kirchler, E. (2006), "Personal Loan Users' Mental Integration of Payment and Consumption". *Marketing Letters* 17, 281-294.

Knutson, B., Rick, S., Wimmer, G. E., Prelec, D. and Loewenstein, G. (2007), "Neural Predictors of Purchases". *Neuron* 53, 147-156.

Lusardi, A. and Mitchell, O. S. (2007a), "Baby Boomer Retirement Security: The Roles of Planning, Financial Literacy, and Housing Wealth". *Journal of Monetary Economics* 54, 205-224.

Lusardi, A. and Mitchell, O. S. (2007b), "Financial Literacy and Retirement Preparedness. Evidence and Implications for Financial Education". *Business Economics* 42(1), 35–44.

Miles, D. (2003), "The UK Mortgage Market: Taking a Longer-Term View". Interim Report to HM Treasury.

Modigliani, F. and Miller, M. H. (1958), "The Cost of Capital, Corporation Finance and the Theory of Investment". *American Economic Review* 48(3), 261-297.

Poterba, J. M. (1984), "Tax Subsidies to Owner-Occupied Housing: An Asset-Market Approach". *Quarterly Journal of Economics* 99(4), 729-752.

Prelec, D. and Loewenstein, G. (1998), "The Red and the Black: Mental Accounting of Savings and Debt". *Marketing Science* 17(1), 4-28.

Prelec, D. and Simester, D. (2001), "Always Leave Home Without It: A Further Investigation of the Credit-Card Effect on Willingness to Pay". *Marketing Letters* 12(1), 5-12.

Puviani, A. (1903), *Teoria della Illusione Finanziaria*. Palermo.

Regeringens proposition 2006/07:1. *Budgetpropositionen för 2007.*

Russell, T. and Thaler, R. H. (1985), "The Relevance of Quasi Rationality in Competitive Markets". *American Economic Review* 75(5), 1071-1082.

Samuelson, W. and Zeckhauser, R. (1988), "Status Quo Bias in Decision Making". *Journal of Risk and Uncertainty* 1, 7-59.

Shafir, E., Diamond, P. and Tversky, A. (1997), "Money Illusion". *Quarterly Journal of Economics* 62(2), 341-374.

Shiller, R. (2005), *Irrational Exuberance.* 2nd edition. Princeton University Press.

Soman, D. (2001), "Effects of Payment Mechanism on Spending Behavior: The Role of Rehearsal and Immediacy of Payments". *Journal of Consumer Research* 27, 460-474.

Soman, D. (2003), "The Effects of Payment Transparency on Consumption: Quasi-Experiments from the Field". *Marketing Letters* 14(3), 173-183.

Stango, V. and Zinman, J. (2007), "Fuzzy Math, Disclosure Regulation, and Credit Market Outcomes". Dartmouth College Working Paper.

Stango, V. and Zinman, J. (2008), "Exponential Growth Bias and Household Finance". Forthcoming in *Journal of Finance.*

Thaler, R. H. (1980), "Toward a Positive Theory of Consumer Choice". *Journal of Economic Behavior and Organization* 1, 39-60.

Thaler, R. H. (1985), "Mental Accounting and Consumer Choice". *Marketing Science* 4(3), 199-214.

Thaler, R. H. and Sunstein, C. R. (2003), "Libertarian Paternalism". *American Economic Review, Papers and Proceedings* 93(2), 175-179.

Thaler, R. H. and Sunstein, C. R. (2008), *Nudge.* New Haven: Yale University Press.

Zellermayer, O. (1996), *The Pain of Paying.* Ph.D. Dissertation, Carnegie Mellon University.

Appendix A

Proof of Proposition 1

Denote

$$x_j^h = P_l - P_h + (E_h - E_l)rt - (E_h^2 - E_l^2)c_j \tag{A.1}$$

The probability that an agent of type $j = 1$ will prefer a high equity apartment over a low equity apartment is the sum of two terms:

$$P_1^h \equiv P\{v_i^a - (P_h) + E_h b_1 > v_i^{a'} - (P_l) + E_l b_1\} \times P\{v_i^a > v^0\} \tag{A.2}$$

The probability that the agent will stay in their current E_h apartment is

$$P_1^0 \equiv P\{v^0 - (P_h) + E_h b_1 > v_i^{a'} - (P_l) + E_l b_1\} \times P\left\{v_i^a < v^0\right\} \tag{A.3}$$

Denote $\Pi_1^h = P_1^h + P_1^0$.

Therefore, with the above notations

$$P_1^h = \frac{1}{2}P\{x_1^h + v_i^a > v_i^{a'}\} = \frac{1}{2}\int_{-V}^{V}\frac{1}{2V}\frac{V + x_1^h + v}{2V}dv \tag{A.4}$$

$$= \frac{1}{2}\Big(\frac{1}{2} + \frac{1}{2V}\frac{1}{4V}(x_1^h + v)^2|_{-V}^{V}\Big) = \frac{1}{2}\big(\frac{x_1^h}{2V} + 0.5\big) \tag{A.5}$$

It is simple to show that the second probability is equal to the first, so $0.5\Pi_1^h = P_1^h = P_1^0$

Thus, the equilibrium conditions will solve

$$\alpha\Pi_1^h + (1-\alpha)\Pi_2^h = \alpha \tag{A.6}$$

in the market for high equity apartments and

$$\alpha\Pi_1^l + (1-\alpha)\Pi_2^l = 1 - \alpha \tag{A.7}$$

in the market for low equity apartments. It follows that

$$\alpha(x_1/2V + 0.5) + (1-\alpha)(x_2/2V + 0.5) = \alpha \tag{A.8}$$

Noting that $\Pi_j^h = 1 - \Pi_j^l$, we get that the market for low equity apartments clears as well.

Plugging in the values for x and multiplying both sides by V we get

$$P_l - P_h + (E_h - E_l)(rt - (1-\alpha)c_2(1-r)\tau(E_h + E_l)) + V = 2\alpha V \tag{A.9}$$

Similarly, for low equity apartments we get

$$P_h - P_l + (E_l - E_h)(rt - (1-\alpha)c_2(1-r)\tau(E_h + E_l)) + V = 2(1-\alpha)V \tag{A.10}$$

Proof of Proposition 3 The probability that an agent of type $j = 1$ stays in their current apartment is

$$P_1^0 = \frac{V + x_1^h}{4V} \tag{A.11}$$

For an for an agent of type $j = 2$ the probability is

$$P_2^0 = 0.5\big(1 - \frac{V + x_2^h}{2V}\big) \tag{A.12}$$

Plugging in x_1^h for $P_l - P_h$ we get $x_1^h = 2V\alpha - V + (1-\alpha)c(1-r)t(E_h^2 - E_l^2)$ and $x_2^h = 2V\alpha - V - \alpha c(1-r)t(E_h^2 - E_l^2)$, implying that for $j = 1$,

$$P_1^0 = \frac{V + x_1^h}{4V} = \frac{2V\alpha + (1-\alpha)c(1-r)\tau(E_h^2 - E_l^2)}{4V} \tag{A.13}$$

and for $j = 2$,

$$P_2^0 = 0.5\big(1 - \frac{V + x_2^h}{2V}\big) = \frac{2V(1-\alpha) + \alpha c(1-r)\tau(E_h^2 - E_l^2)}{4V} \tag{A.14}$$

Note that

$$\frac{\partial P_1^0}{\partial c} = \frac{(1-\alpha)(1-r)\tau(E_h^2 - E_l^2)}{4V} \geq \frac{\partial P_2^0}{\partial c} = \frac{\alpha(1-r)\tau(E_h^2 - E_l^2)}{4V} \tag{A.15}$$

as soon as $\alpha \leq 0.5$.

CHAPTER 3

Lady and the Trump: Status and Wealth in the Marriage Market

with Anna Dreber[1]

Abstract. We examine a relatively neglected aspect of intergenerational transmission of economic standing, namely culturally determined status markers and their valuation in the marriage market. We take nobility to be such a status marker. Using data on Swedish marriages, we test the hypothesis that nobility have a greater probability of marrying "up" in terms of wealth. We find a large and statistically significant positive effect for nobility. This finding has implications for the intergenerational transmission of inequality, and for the longevity of the institution of nobility itself.

1. Introduction

If wealth begets wealth, dynasties may endure even in otherwise meritocratic societies. And if an indicator of past wealth begets wealth, this too will reinforce the persistence of economic standing across generations. Such indicators typically bestow prestige and contribute to high status. Nobility, a culturally determined (i.e., non-genetic) hereditary status marker, might act as such an indicator, and thereby serve as a vehicle for the cultural transmission of economic standing. A wide range of channels besides the traditional economic variables have previously been found to be incorporated in the intergenerational transmission of economic standing. For example, heritability of physical traits, such as cognitive ability and health, and physical appearances, such as height, attractiveness, and race, have all been found to affect economic outcomes (see, e.g., Hamermesh and Biddle, 1994; Bowles and Gintis, 2002; Tao, 2008). While

[1] This paper originally appeared as the lead article in *Kyklos* 62 (2009). We thank Tibor Antal, Sam Bowles, David Cesarini, Robin Douhan, Magnus Henrekson, Henrik Horn, Yannis Ioannides, Magnus Johannesson, Erik Lindqvist, Karen Norberg, Martin Nowak, Thomas Pfeiffer, Mark Voorneveld, Daniel Waldenström, David Yermack, seminar participants at the Research Institute of Industrial Economics, Kyklos editor René Frey and two anonymous referees for helpful comments. Johan Almenberg thanks the Jacob Wallenberg Foundation and the Torsten and Ragnar Söderberg Foundations for financial support, and Johan Almenberg and Anna Dreber thank the Jan Wallander and Tom Hedelius Foundation for financial support. We thank Johanna Chamberlain for research assistance.

cultural transmission of economic standing clearly also plays a part, this vehicle has so far remained relatively unexplored.

Nobility as an institution is an anachronism: it is a traditional term for the highest social class in some pre-modern societies. Status, however, plays an important role in most societies and in most times, making individuals allocate valuable resources to status-enhancing activities. This suggests that even though nobility no longer entails formal privileges, it may continue to be coveted as a status marker.

Anecdotal evidence suggests that nobility has retained its allure in the modern age. John von Neumann, the mathematician who pioneered the expected utility framework and laid the foundations of modern game theory, was the son of Max Neumann, who in 1913 purchased a claim to Austrian nobility. Max Neumann thereby acquired the right for his offspring (but not for himself) to call themselves von Neumann, a right of which his son John made good use.[2] Another example is German industrialist Heinrich Thyssen, who married the Hungarian baroness Margit Bornemisza in 1906. Thyssen, a commoner, had his father-in-law adopt him, and since the Baron had no male heirs, Emperor Franz Joseph I bestowed on Thyssen and his descendants the right to adopt the Bornemisza name, coat of arms, and title of Baron. More recently, a scandal erupted in the UK in 2006 following indications of a correlation between individuals making large loans on favorable terms to the Labour party, and subsequent nominations for peerage.[3]

We examine the relative performance of nobility in the marriage market. In doing so, we seek to fuse research on status with research on mating patterns. Our enquiry rests on two implicit assumptions: (1) nobility bestows status on the beholder; and (2) individuals get utility from status, either directly or indirectly. Nobility is typically not traded in open markets, making it difficult to observe its price. We argue that the marriage market might serve as an informal conduit for such transactions. If nobility is a vehicle for the intergenerational transmission of economic standing, the marriage market is the mechanism for this transmission. If such indirect trade occurs, we ought to be able to observe the valuation of nobility in this market.

The Swedish marriage market provides us with an opportunity to estimate the valuation of nobility. In Sweden, nobility as an institution originates from the Alsnö Rules of 1280, which granted landowners exemption from taxation in exchange for supplying the monarch with cavalry troops (vassalage). During the Middle Ages, the link between vassalage and membership of the nobility became weaker. Increasingly, noble titles came to be handed out at the monarch's discretion, and it is these titles

[2] By contrast, von Neumann's contemporary Friedrich von Hayek was *bona fide* nobility, but preferred to omit the von and simply call himself Friedrich Hayek.

[3] These peerages are not hereditary.

that have come to constitute the institution of nobility in its present form. The vast majority of the nobility was created in the period 1611-1718, a period of more or less continuous warfare. The last time Sweden entered into war was 1814. In the ensuing peacetime, the creation of nobility declined rapidly. During the 19th century, the monarch's right to hand out noble titles was increasingly questioned. The nobility lost most of their formal privileges in the reform of 1809, and their political influence was greatly reduced through the reform of 1866, in which the House of Nobility was stripped of its role as upper chamber of Parliament. In 1975, the monarch's right to hand out noble titles was formally revoked.[4] Today, Swedish nobility enjoy no formal privileges. Swedish law does not permit transferring ownership of a claim to nobility in an open market. Marriage is the only remaining conduit for those seeking to join the nobility. In sum, nobility is an asset that conveys no material privilege and cannot be traded in an open market.

An old Swedish custom enables us to identify members of the nobility in our marriage data. Beginning in the mid-16th century, it became customary for newly created nobility to take a new, distinct name upon becoming part of the nobility, often using a familiar set of prefixes and suffixes.[5] Moreover, Swedish law awards intellectual property rights to surnames in direct relation to how distinct they are, i.e., in inverse proportion to the number of families sharing the name. As a consequence, noble names enjoy particularly strong protection and are easily identifiable. The House of Nobility in Stockholm publishes an annual directory of the members of the approximately 600 remaining noble families. Combining the records of the House of Nobility with the Total Population Register compiled by Statistics Sweden enabled us to generate a unique data set, consisting of repeated cross-sections of all marriages in Sweden in 1985, 1990, 1995, 2000 and 2004. In addition to information about the age, education, income and wealth of both spouses, the data also contains an indicator showing if an individual's surname denotes nobility.

Our data set provides an opportunity to examine whether people are willing to trade wealth for status, by testing the hypothesis that the probability of hypergamy (marrying "up") in terms of wealth increases when an individual belongs to the nobility. If nobility bestows status, and if individuals value both status and material consumption, we would expect an individual belonging to the nobility to attain a premium in

[4] At that point, this right had not been exercised in a long time: the last individual to join the ranks of nobility was the explorer Sven Hedin, in 1902.

[5] A few noble families bear common names (i.e., names that are not distinct). If there are individuals in our data set that have mistakenly been identified as nobility, this might result in a slight downward bias in our estimates.

the marriage market compared to a non-noble individual with otherwise identical characteristics. Such mating patterns would be consistent with complementarities between wealth and status.

We also test the auxiliary hypothesis that male nobility is valued more than female nobility. To the extent that individuals care about their offspring, they will attach greater value to a status marker if it can be passed on to their children. Swedish nobility is hereditary on the male side only. Conditional on a continued male lineage, noble status can be thought of as an asset that continues to pay dividends indefinitely. Even with intergenerational discounting, we would expect such an asset to attain a higher price than the non-hereditary equivalent, which can be thought of as an asset paying a dividend only in the current period.

Our main finding is a significant increase in the probability of hypergamy in wealth for members of the nobility, controlling for own wealth and other covariates. This nobility premium is sizeable. The wealth distribution in our data is heavily concentrated in the lowest wealth bracket, resulting in a low baseline probability of marrying up. Given this, the observed nobility premium must be considered large, increasing the probability of marrying up by about 40 percent. The effect is statistically significant and robust to a number of different measures of hypergamy.

We find less support for the auxiliary hypothesis that male nobility attain a higher premium than female nobility. The interaction term for male sex and nobility has the expected (positive) sign, but is not statistically significant. In other words, we are unable to reject the hypothesis that the premium is of equal magnitude for male and female nobility.

In sum, our main finding is consistent with the hypothesis that nobility attain a premium on the marriage market, indicating that mate preferences are status sensitive. Our results are highly robust to different definitions of hypergamy. This suggests that the cultural transmission of economic standing, channeled through nobility and other status markers, should be taken into account in an analysis of the intergenerational transmission of inequality.

Our paper is organized as follows. In section II, we outline some relevant existing research on status and marriage. In section III, we describe our marriage data. In section IV we present our econometric model and report the regression estimates based on the marriage data. We also summarize our robustness checks. We conclude in section V, where we discuss some implications of our results and suggest directions for future research.

2. Related literature

Our analysis draws on two strands of economic research: on the one hand, literature looking at the role of status in the economy, and on the other hand, the economic analysis of the "marriage market", i.e., the matching of brides and grooms. In this section we outline some of the more relevant literature in these areas.

2.1. Status. That the concern for relative position plays an important role in social interactions is by no means a recent insight in economics. Adam Smith devoted a considerable part of *The Theory of Moral Sentiments* to a discussion about the link between wealth and social esteem (Smith, 2000, see in particular chapters 2 and 3). Veblen (1899, pp. 25-26) famously argued that the pursuit of relative position is the main driving force of (conspicuous) material consumption:

> The motive that lies at the root of ownership is emulation; and the same motive of emulation continues active in the further development of the institution to which it has given rise and in the development of all those features of the social structure which this institution of ownership touches. The possession of wealth confers honour; it is an invidious distinction.

This raises the question of whether relative position is pursued as an end in itself. Frank (1999) argues that humans may well be hardwired to seek gratification from moving up in the social hierarchy. He points to research showing that relative position is correlated with serotonin levels in non-human primates (McGuire, Raleigh and Brammer, 1982, Raleigh and McGuire, 1994).

If individuals derive utility from status in itself, independently of material benefits associated with it, a utility function that incorporates relative position as one of its arguments might do a better job of explaining individual choices (see, e.g., Frank, 1985; Huck and Müller, 2000). The idea of status in the utility function has recently been picked up by Becker, Murphy and Werning (2005, p. 283), who note that "[w]hen status is important, individuals would be willing to pay a lot in time, effort, and money for sufficiently high status." Utility derived from relative position in itself has also been analyzed in theoretical work on incentives (see, e.g., Moldovanu, Sela and Shi, 2007). A closely related discussion can be found in Frey (2006), who analyzes the supply of, and demand for, awards such as medals, orders and decorations, and emphasizes their role as status markers. Another related strand of research takes identity, of which status may be an important determinant, as an argument of the utility function (Akerlof and Kranton, 2000).

To the extent that other individuals attach a value to relative position, status markers may have economic value even to individuals who do not care about relative

position in itself. In a laboratory experiment, Ball *et al.* (2001) find that economic outcomes are affected by the status of participants in a market. Individuals who were assigned high status attained higher prices as sellers, and lower prices as buyers, than individuals who had been assigned low status. This effect prevailed even when the assignment of status was entirely random, and the randomization process was common knowledge. Such behavioral patterns would allow individuals to reap direct economic benefits from status markers. Hence they may covet these without necessarily getting any utility from the status marker in itself. We return to this important qualifier in section 2.3.

2.2. The marriage market. We hypothesize that individuals with wealth but no status are likely to be matched in marriage with individuals with status but less wealth. Such a mating pattern is assortative, in the sense that individuals are sorted non-randomly into matched pairs on the basis of observable characteristics. Most animals engage in assortative mating. The most commonly observed pattern is positive assortative mating, meaning that a member of a species is matched with another member of that species who is similar with regard to a certain trait. Compared to less stratified mating patterns, positive assortative mating reinforces differences in the endowments of biological, economic and cultural assets in the population over time (Fernández and Rogerson, 2001; Fernández, Guner and Rogerson, 2005). Under certain assumptions, such mating patterns can preserve heterogeneity in a population indefinitely (Bisin and Verdier, 2000).

The seminal microeconomic model of assortative mating in the marriage market is Becker (1973, see also Becker, 1991). This model, originally introduced by Koopmans and Beckmann (1957) to describe the assignment of plants to locations, generates a pairwise assignment of elements of one set to elements of another set. The elements of a set differ in a single trait. In Koopmans and Beckmann (1957), an optimal sorting is defined as a sorting that is in the core, in the sense that there is no other coalition outside the core in which both parties could be made better off by an alternative sorting. A key result in this model is that such a sorting necessarily produces the greatest aggregate output (summing across all matches), though it does not necessarily include the largest output element (the matching with the largest output).

Becker's adaptation of the Koopmans and Beckmann model to the marriage market hinges on his characterization of marriage as an economic institution. Central to his analysis is the idea of a marriage production function, which takes the endowments (in a wide sense) of both spouses as its inputs, and produces a joint output that is consumed by the household. This joint output consists of pecuniary rewards from the labor market, non-pecuniary production taking place at home, and leisure. The

model is unitary, in the sense that the household is assumed to maximize this output according to a single utility function, which does not require explicit modelling of the bargaining for resources within the household.

In Becker's (1973) model, individuals differ in a single trait, common to both men and women, and a marital output function that is increasing in this trait. The model predicts positive assortative mating in this trait if an increase in the trait for both partners has an effect on marital output that is greater than the sum of the partial effects of increasing one while holding the other constant. In other words, increasing the trait in one partner raises the marginal effect of the other partner's endowment of the trait on marital output. In economic terms, there is *complementarity* between the two partners' endowments of the trait. In the presence of such complementarity, it can be shown that positive assortative mating produces the greatest aggregate output over all marriages.[6]

Positive assortative mating has been reported in a variety of traits, for both humans and animals. Almenberg and Dreber (2009) show that Swedish marriage data indicate positive assortative mating within both status and wealth.[7]

In the case of mixed matches between wealthy individuals and individuals with high status, however, the assortative mating is between traits rather than within a trait: individuals with high wealth but low status sorting with individuals with low wealth but high status. To the best of our knowledge, there have been no previous empirical economic examinations of assortative mating between status and wealth in the marriage market.

2.3. Status and marriage. Economic theory blending the themes of status and marriage is scarce; these areas have typically been examined separately. A notable exception is Mailath and Postlewaite (2006, henceforth M&P), who introduce the notion of a "social asset", being an asset that derives some or all of its value from social institutions. They distinguish between the fundamental and intrinsic values of an asset. A status marker that has no fundamental value may still have intrinsic value if it raises expected future income.

While typically non-transferable within a generation, M&P suggest that certain hereditary traits might have such instrumental value because parents typically care about not only their own consumption but also the consumption of their offspring: while hereditary traits cannot be traded contemporaneously, individuals who don't

[6] A proof is given in Becker (1991, p. 130). Note that the model is frictionless, in the sense that the optimal matching is assumed to come about of itself, without any reference to how the process of sorting actually takes place.

[7] There is a large literature on positive assortative mating as well as mate preferences in general. See, e.g., the references in Almenberg and Dreber (2009)

possess the trait can mate with individuals who do, in the hope that their offspring will inherit it.

The value of hereditary social assets may in part be due to the difficulty of insuring future generations against consumption risk: while any generation may squander a family's economic resources, the subsequent generations can still be endowed with a single asset that cannot be relinquished, and yet raises their expected consumption – namely, a trait that bestows status. In other words, each generation of a lineage can extract the flow value of the asset, but cannot extract the capitalized present value of future flows.

The valuation of such social assets may be self-fulfilling. If everybody else attaches value to an asset, it may become covetable also for individuals to whom it has no intrinsic value. Agents in the M&P model differ in terms of income (which is either high or low) and a binary trait. Income is assumed to be non-storable, so that parents cannot transfer consumption to their offspring. If individuals in possession of the trait are *ceteris paribus* more desirable partners, they will have a greater chance of marrying "up" in wealth. In this case, acquiring the trait through marriage becomes a means for the parents to insure against some of the consumption risk of their offspring. In other words, the best response of an individual in a society where others seek to marry individuals with the coveted trait may be to also seek to marry individuals with that trait, thereby raising the chances of their offspring having the trait and thereby higher expected income. Hence, in equilibrium the desirability of the trait can be self-fulfilling.[8]

The particular characteristics of the attribute are irrelevant: "Any heritable attribute might serve as a social asset in this way" (Mailath and Postlewaite, 2006, p. 1059). Although M&P do not mention hereditary nobility, this institution fits the prescribed mechanism well. It is widely accepted that broad measures of ability such as IQ are mean-reverting. Noble titles were typically awarded in reward for distinguished service to the monarch. Such services often also resulted in pecuniary rewards, such as land grants. Family fortunes, however, can be lost in the course of a single generation, whereas hereditary nobility is inalienable: it cannot be sold since it cannot be transferred. For an individual with high ability and high income, nobility may have offered a means to insure against lower ability in a subsequent generation, and the associated risk of an inferior economic outcome for the individual's offspring.

[8] The offspring have higher expected income for two reasons: income may be correlated with the trait, and possession of the trait raises the likelihood of marrying a high income individual who in turns wants to insure the consumption risk of his/her offspring. The first mechanism is not a necessary requirement.

3. Marriage data

To shed more light on the role of nobility in marriage markets, we use a repeated cross-section of all marriages in Sweden during 1985, 1990, 1995, 2000 and 2004, in total 195,405 marriages.[9] We are interested in the individual probability of marrying up in wealth, hence each spouse is treated as a separate observation, giving us a sample of 380,810 observations. The data are drawn from the Total Population Register, contain every registered marriage during the years in question, and were compiled by Statistics Sweden on our behalf. The data contain information on a number of characteristics of bride and groom, including age, income, net wealth and level of education, all measured *at the time of marriage*. The data set also contains a dummy variable that takes the value 1 if an individual's surname denotes nobility, and 0 otherwise. A set of names belonging to the remaining families of the nobility was provided by the House of Nobility in Stockholm and used to generate this indicator. In the data set, 1,782 individuals belong to the nobility according to this indicator, equivalent to a bit less than 0.5 percent of the sample.[10]

The control variables are defined in terms of brackets. In part this reflects the requirement from Statistics Sweden to protect individual integrity, given the small number of nobility in our sample. There are five (annual) income categories: [0–121,999], [122,000–199,999], [200,000–299,999], [300,000–499,999], and [500,000–].[11] Age is in the following brackets: (–24], [25–29], [30–34], [35–39], [40–44], [45–54], or [55–). The data on education places each individual in one of four categories, corresponding to pre-high school, high school, less than three years of tertiary education, and more than three years of tertiary education. Individual wealth belongs to one of four categories: [0–199,999], [200,000–499,999], [500,000–1,499,999], and [1,500,000–]. There is a high concentration of individuals in the lowest wealth bracket. For each of these covariates, we construct a set of dummy variables corresponding to the aforementioned categories. In the regressions, the lowest bracket is the baseline for each categorical variable.

The wealth tax that was effective in Sweden until 2007 provided strong incentives for tax avoidance. Assets not taxed at all or entered in tax returns at levels below

[9] At the time of writing, the data for 2005 was not yet available, so we chose the closest available data.

[10] This is slightly larger than the proportion of nobility in the general population (about 0.3%). This could be caused by nobility having an above average propensity to marry. Bearing this in mind, we can only make inferences regarding the individuals that actually marry. It could also be caused by sampling error, due to some noble names not being distinct. In this case our estimates will be biased downward, thus underestimating the actual nobility premium.

[11] All amounts reported are nominal and measured in SEK. 1 SEK $\approx$ 0.1 USD, adjusted for purchasing power.

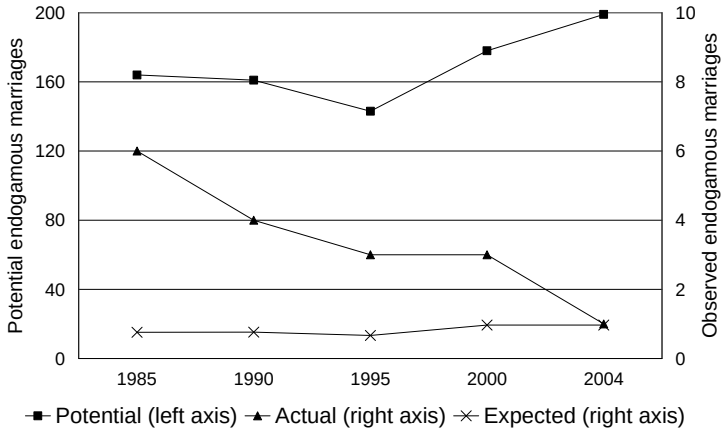
market value included real estate, land holdings, art and antiques, and shares of small-cap firms on the Stockholm stock exchange (not on the A-list). Given that there have been numerous ways of reducing taxable wealth in Sweden, it is unlikely that these figures show the individuals' full wealth. This reduces the efficiency of our estimator. In addition, it is a potential source of unobserved heterogeneity in the sample. Let W be the true wealth, and $(1-\alpha)W$ the observed wealth of individual i, with $0 < \alpha < 1$. If α is positively correlated with belonging to the nobility, our estimate of the nobility premium will be biased upwards. This could make what might in fact be positive assortative mating look like hypergamy. Note that even if α is correlated with nobility, our estimates of the difference in the nobility coefficient between male and female nobility will not be biased unless the correlation itself differs systematically between the sexes. We discuss unobserved heterogeneity more extensively in the next section.

4. Hypothesis testing

4.1. From marrying "in" to marrying "up". Traditionally, nobility were expected to marry within their own ranks. The marriage market data indicate that this is changing. Endogamy (within-group marriage) still occurs: the odds of a spouse belonging to the nobility are more than twice as high if an individual also belongs to the nobility (logit regression, $\beta_{Nobility}$: 1.20; p-value < 0.001). The coefficient on the nobility dummy is large compared to the coefficients on age, income or wealth. Yet, the tradition of endogamous marriage seems to be coming to an end. The great majority (98%) of the nobility in our sample are not married to other nobility, and the number of marriages in which both spouses belong to the nobility declines monotonically over time. Figure 1, below, shows the actual number of endogamous marriages among Swedish nobility observed in our sample, by year. This is plotted against the potential number of such marriages, defined (somewhat arbitrarily) as the minimum of male/female nobility getting married in that year. The third series shows the expected number of such marriages, if nobility were randomly assigned to males and females in our sample. There are two salient features of Figure 1. First, the frequency of endogamous marriage has been steadily declining, despite a more or less constant number of nobility in the marriage market. Second, the rate of endogamous marriage has been converging toward, and recently reached, the level that we would expect to observe with random matching.

Figure 1

Endogamous Marriages among Swedish Nobility, 1985-2004



In other words, endogamous marriage within the nobility, once considered the norm, has become an unusual occurrence.[12] For the most recent observations in our data, such marriages occur at a rate no greater than what we would observe under random matching. One interpretation of this pattern is that nobility, in response to modernizing reforms that curtail their economic influence, turn to exogamy as a dominant strategy for securing access to resources. This process may have gone on for some time: for example, strategic marriages between British nobility, rich in symbolic capital but cash poor, with the daughters of American industrialists in the late 19th and early 20th century have been documented by historians (see, e.g., Cannadine, 1990).

In Sweden, the waning influence of the nobility in the 20th century was reflected in a marked decline in their statistical overrepresentation on company boards, in the foreign service, and other prestigious positions (Rundblad, 1999). In 1968, individuals belonging to the nobility constituted 12 percent of all board members of the 50 largest companies on the Stockholm stock exchange. By 1998, this had declined to 4 percent.[13] In the foreign service, nobility constituted 26 percent of all ambassadors and consul generals in 1968. This had declined to 8 percent by 1998 (Rundblad 1999). When the formal privileges of nobility were curtailed, marrying up in wealth may have gained importance as a means of securing access to material resources.[14]

[12] By contrast, Banerjee et al. (2008) report that within-caste marriage is still prevalent in India.

[13] As measured by turnover (Rundblad 1999).

[14] Note that from a long-term perspective, this loss of privileges coincided with the emergence of the Swedish welfare state and a reduction in income inequality. Frey (2006) argues that awards should play a more important role as status markers in countries with low income dispersion and less pervasive market ideology. To the extent that this analysis carries over to nobility, an inherited award,

4.2. Regression analysis: is there a nobility premium? Having noted that nobility are not marrying within their group to any great extent, we now turn our attention to whether they are marrying up in wealth. According to our hypothesis, wealthy individuals will covet status markers in general and hereditary, inalienable status marker in particular, resulting in a higher probability of marrying up in wealth for individuals belonging to the nobility. To test this hypothesis, we construct an indicator variable that takes on the value 1 if individual i marries into a higher wealth bracket, and 0 otherwise. To examine the robustness of our results, we test a number of different specifications of this indicator, including the transition probabilities between specific wealth brackets. Our results are broadly robust to such modifications.

We use the logistic binary response model (logit regression) to estimate the probability of an individual marrying a spouse in a higher wealth bracket. The logistic functional form generates estimates that are bounded on the unit interval, and hence has an intuitive appeal when the dependent variable is a binary outcome.

We regress a dummy variable that indicates marrying up on an individual's age, education, income, wealth, and sex, as well as whether the individual belongs to the nobility.[15] We have no reason to believe that the parameters of the true model are the same for both sexes. For this reason, we run separate regressions for women and men.[16] Since we are also interested in how the nobility coefficient differs between the sexes, we also run a joint regression where we allow all coefficients to differ between the sexes. We include a dummy variable indicating whether the individual is male, allowing for different baseline probabilities for men and women, and interaction terms between the male dummy and the independent variables, including the nobility dummy, thus allowing for the slope coefficients to differ between the sexes.

The joint regression offers a statistical measure of how the nobility coefficient differs between the sexes. A Wald test rejects that these interaction terms are jointly insignificant (p-value < 0.0001). In each regression, we also include a full set of controls for the spouse, to make sure that we are not confounding an increased probability of marrying up in wealth with an increased probability of marrying somebody with higher age, income or education. Note that in order to run a joint regression, we rearrange the observations so that each individual appears twice, once as individual i and once

these developments could in fact have raised the value of nobility in the marriage market at the same time as nobility were losing their explicit privileges.

[15] Since individuals in the highest wealth bracket cannot marry up, we exclude these observations from the analysis. In our main regression, we do not control for the spouse belonging to the nobility. Our results are robust to the inclusion of this control. They are also robust to omitting all spouse controls.

[16] Our approach is similar in spirit to that of Elmslie and Tebaldi (2008) who examine socioeconomic determinants of infidelity.

as a spouse. This allows us to estimate the probability on the whole population of individuals in the sample. For the regression on the joint sample, where we allow the coefficients to differ between the sexes, we can write our regression specification as a logistic probability model of the form

$$P(Y_{ij} = 1|X_{ij}) = 1/(1 + e^{-(\alpha_j + x_{ij}\beta_j + \delta_j Nobility)}) \tag{4.1}$$

where $P(Yij = 1|X_{ij})$ denotes the probability of marrying up for an individual i of sex j, conditional on the covariates. Nobility is a dummy variable that takes on the value 1 if individual i belongs to the nobility, and is the vector of controls mentioned above. The subscript j denotes gender. Table 1 summarizes the results from the regressions, for both sexes separately (columns 1 and 2), and for the joint sample (column 3).

Table 1

Dependent Variable: Marrying "Up" in Wealth

	Women	Men	All
	(1)	(2)	(3)
Nobility	0.390	0.434	0.382
	(0.021)**	(0.031)**	(0.024)**
Male*Nobility			0.055
			-0.834
Additional Controls	Yes	Yes	Yes
Interaction terms (*male)	No	No	Yes
Male			-0.473
			(0.000)***
Constant	-5.129	-5.517	-4.919
	(0.000)***	(0.000)***	(0.000)***
N	171187	170464	341651
R^2/Pseudo-R^2	0.141	0.141	0.146

Robust standard errors. *p* -values in parentheses

* significant at 10%; ** significant at 5%; *** significant at 1%

The coefficient for nobility is positive and statistically significant in all specifications. Columns 1 and 2 show that both male and female nobility have higher probabilities of marrying up in wealth, when controlling for the available covariates. The coefficient is larger for noble men than for noble women. Column 3 shows the same thing, but the difference is not statistically significant. We rely on the specification in column (3) of Table 1 as our main model, and interpret our results below.

The observed nobility premium is sizeable. The wealth distribution in our sample is heavily concentrated in the lowest bracket. When an individual marries up, his/her spouse by definition marries down. Hence, the largest potential number of marriages in which one individual marries up in wealth equals the number of individuals in the

second, third and fourth wealth brackets, about 3.2% of the sample. As a result, marrying up in wealth is a rare event, achieved by only 2.6% of the individuals in our sample. For nobility, however, this frequency is 4.4%. In essence, our regressions show is that including a number of relevant controls does not diminish this nobility premium. In the main regression (Table 1, column 3), the estimated nobility premium is an almost 40 percent increase in the odds of marrying into a higher wealth bracket.

While the interaction term between nobility and being male has the expected (positive) sign in the logit regression in column 3, it is not statistically significant. This implies that we cannot reject that the increase in probability is of the same magnitude for noble men and noble women. Thus, we do not find conclusive support for our auxiliary hypothesis of a gender gap in the nobility premium.[17]

Nobility might be correlated with variables that we are unable to control for in the sample, for example human and social capital, or even physical appearance.[18] If such unobserved heterogeneity is a direct consequence of nobility, this does not undermine the validity of our findings. It is important to recognize, however, that there might be other unobserved variables that could arguably affect our results. This is a further reason for interpreting our results with some caution. Nobility are slightly over-represented among individuals that marry, implying that there might be some selection bias. Moreover, if the fraction of an individual's wealth that goes unreported is positively correlated with belonging to the nobility, then our estimates will be biased upward. Nobility are "to the right" of the general sample in terms of the wealth distribution. Nobility may have wealthier parents than other individuals in the same wealth bracket. An individual marrying a member of the nobility with no wealth could be expecting future wealth through inheritance. Another possible omitted variable would be that past wealth is positively correlated with residing in affluent neighborhoods, and residing in affluent neighborhoods may be positively correlated with the probability of marrying "up".

A more sophisticated modeling approach to the marriage market would add additional complexity to the interpretation of our results. Hypothetically, nobility could be more prevalent in marriage markets in which population densities, the fraction of singles, and sex ratios among singles differ from the population average. In a search model of the marriage market, these factors would be expected to influence the reservation price at which a match is made (Drewianka, 2003). Controlling for such factors

[17] We also estimate a linear (OLS) model using the same specification. With this functional form, the interaction term between male sex and nobility has the wrong sign, and is still not statistically significant. Because the R^2 of the linear model is considerably lower than for the logistic model, we focus mainly on the latter. Estimates for the linear model can be provided upon request.

[18] For an economic perspective on social capital, see, e.g., Sabatini (2008).

is beyond the scope of our data set, but would constitute an interesting avenue for further research.

While each of these concerns is valid, they affect only the interpretation of our findings, and not their implications. Regardless of whether nobility marry up in wealth because their nobility is a covetable status marker or because they live in affluent neighborhoods etc., the consequence of this hypergamy is that economic resources are channeled toward nobility at a time when they have relinquished all explicit economic and political privileges. To the extent that access to resources is linked to reproductive success, this may have considerable implications for the longevity of nobility as an institution and for the intergenerational transmission of inequality.

4.3. Robustness checks. We conduct a number of robustness checks to test the validity of our results. Our conclusion, based on these checks, is that our results are robust. A non-technical summary of the robustness checks follows below.

Our first robustness check consists of examining whether our main finding, the nobility premium, is consistent with alternative measures of hypergamy. The measure we use for our main regression, reported in Table 1 above, is the probability of marrying an individual in a higher wealth category. Let x_i , y_i denote the wealth bracket of individual i and individual i's spouse. Our main model corresponds to estimating the probability of $y_i > x_i$, conditional on the covariates. To check the robustness of our findings, we use the same model specification, except that we use three other measures of hypergamy: (1) $x_i = 1$, $y_i \geq 2$; (2) $x_i \leq 2$, $y_i > 2$; and (3) $x_i = 1$, $x_i = 4$. We find evidence of a nobility premium of similar magnitude in all the aforementioned cases, suggesting that our results are not particularly sensitive to the choice of hypergamy measure. Coefficients and p-values for the nobility dummy are shown in Table 2, below.

Table 2

Robustness Checks

	Nobility dummy	
Measure of hypergamy	Coefficient	p-value
$x_i = 1, y_i > 1$	0.499	(0.003)***
$x_i < 3, y_i > 2$	0.400	(0.077)*
$x_i = 1, y_i = 4$	0.862	(0.003)***

Robust standard errors. p-values in parentheses
*significant at 10%; **significant at 5%; ***significant at 1%

A second purpose of these checks is to examine whether our results are driven by differences in the distribution of nobility and non-nobility within wealth brackets.

The wealth distributions of the nobility and of the general sample are different, with the probability mass of the former being "to the right" of the latter, in the sense of stochastic first order dominance. Given the high level of aggregation – only four wealth brackets – differences in the distribution within each bracket might influence our results. If, for example, the average wealth of nobility in the lowest wealth bracket is higher (lower) than average wealth of non-nobility in the same bracket, our estimate for this group would be biased upward (downward). As an additional precaution, we also estimate the probability of marrying "down". We run a regression similar to the one above, except that the dependent variable is a dummy variable indicating whether an individual from the second, third or fourth wealth bracket marries an individual in a lower wealth bracket, $x_i > y_i$.[19] If nobility in the higher wealth brackets on average have higher wealth than other individuals within that bracket, then they should be less likely to marry down. We find no evidence of such an effect (nobility dummy p-value < 0.971). We conclude that our results are not likely to be driven by differences in the distribution of nobility and non-nobility within wealth brackets.

5. Conclusion

We have attempted to fuse the research on status with that on mating patterns, by examining the relative performance of hereditary nobility, a proxy for status, in the marriage market. In Sweden, nobility no longer enjoy formal privileges, yet anecdotal evidence suggests that nobility remains coveted. Since nobility is not traded in open markets, we study the marriage market as a conduit for such transactions.

We find that nobility enjoy a higher probability of marrying up in wealth. Because our sample is heavily concentrated in the lowest wealth bracket, we observe few individuals marrying up in wealth. For the general sample, the rate is 2.6%. For the subsample belonging to the nobility, however, this rate is 4.4%. We have shown that this apparent nobility premium is robust to controlling for a number of relevant variables.

The data also indicate a gender difference, with males, for whom nobility is hereditary, attaining a higher premium. The gender difference, however, is not statistically significant. Taken together, these two results could be explained in several ways. To suggest a few: (1) There is a nobility premium, and it is higher for males. The sample contains enough individuals belonging to the nobility to detect the premium, but not enough to identify the gender difference. (2) There is a nobility premium, and it is

[19] Note that this regression is, in practice, not perfectly symmetrical with a regression looking at the probability of marrying up. The former omits individuals in the lowest wealth bracket, who by definition can't marry down. The latter omits individuals in the highest wealth bracket, who by definition can't marry up.

higher for males. Other gender differences in mate preferences obscure the gender difference in the nobility premium. (3) There is a nobility premium, because individuals care about their own status, but no gender difference, because individuals do not care about their children's status.[20] (4) There is no nobility premium. The observed effect is driven by unobserved heterogeneity.

Our data does not enable us to discern which, if any, of these alternative explanations is the correct one. Hypothesis testing on our sample of observational data rejects the null hypothesis that nobility do not have better chances of marrying up in wealth. The data has clear limitations, and we are not able to rule out concerns about unobserved heterogeneity – as is always the case with observational data. We wish to emphasize that the results presented here are a first step toward a better understanding of the nobility premium, and we encourage other researchers to shed more light on this topic. We will conclude by discussing some possible implications of a nobility premium.

Who marries whom influences future generations to the extent that the characteristics of the parents are passed on to their children, through their genes as well as a shared environment. Who marries whom is therefore an indicator of the distribution of a wide range of characteristics of successive generations. In a similar manner, the distribution of culturally determined status markers in the population may have important effects on economic outcomes.

In an era when their formal privileges have been curtailed, a marriage premium for nobility suggests a positive valuation of the status marker itself. If the symbolic capital of nobility continues to attract a premium in the marriage market, this provides an additional mechanism that reinforces the persistence of social stratification and inequality. Moreover, if status and consumption of other goods are complementary as Becker, Murphy and Werning (2005) suggest, and if status markers are becoming relatively more scarce, the nobility premium might well be increasing over time.[21]

If access to material resources is correlated with greater reproductive success, a marriage market premium would imply that nobility may persist for a long time, even if no new nobility is created. Historically, male nobility, and rich men in general, have reproduced to a greater extent than other men (e.g., Clark, 2007). This is not surprising given that these men tended to be in the upper strata of wealth and status, and had

[20] This would contradict the assumptions of the model in Mailath and Postlewaite (2006). It would also contradict the large literature on bequest motives.

[21] Becker, Murphy and Werning (2005) show how the price of status relative to other consumption goods increases when status markers become relatively more scarce. In a similar vein, Frey (2006) argues that the value of an award is increasing in its scarcity. It is plausible that this applies also to inherited awards.

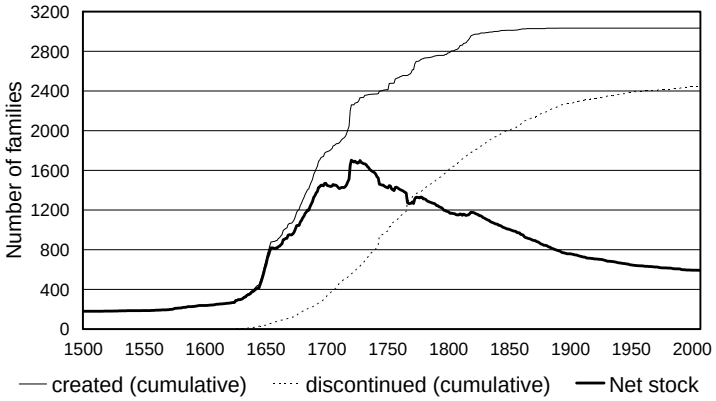
privileges held by neither commoners nor noble women. Wealth and status have both been found to correlate positively with male reproductive success among other groups, such as 19th Century Mormons (Mealy, 1985) and the Ifaluk in Micronesia (Turke and Betzig, 1985).

When nobility is only passed on through male offspring, the perpetuation of nobility requires a continued male lineage. We merged data in Fahlbeck (1899) with 20th Century records from the House of Nobility in order to trace the evolution of the stock of noble families over time.[22] Note that the number of families is the observed quantity, and has not been corrected for population size.[23]

Figure 2 shows the evolution of the stock of noble families since the year 1500. The net stock (thick solid line) is the difference between the number of families that have been elevated to nobility (thin solid line) and the number of noble families whose male lineages have been discontinued (dotted line).

Figure 2

The Evolution of the Stock of Noble Families Since the Year 1500



A rapid decline in the creation rate has caused the stock of nobility to decline monotonically in the last 200 years (i.e., in peacetime). The recent decline is well

[22] These records are as yet incomplete: while it is generally known if a male lineage has been discontinued, it is not always known in which year this happened. In the cases where the decade, but not exact year, was known, we used the middle year of that decade. In the cases where not even the decade was known, the discontinuation year was imputed using linear regression. The discontinuation year was imputed using information on year of creation, monarch at time of creation, and the occupation of the first member of the noble lineage. Approximately 10 percent of all discontinuation dates were imputed in this manner.

[23] The analysis is made slightly more complicated by the fact that the dataset includes families that were raised from the lower ranks of nobility to the higher, titled ranks. This applies to about 20 percent of the sample. The branch that was elevated to a higher rank is counted in the data as the creation of a new noble family, which overlooks the fact that it stems from a male lineage that is already noble. Excluding this subset, however, has negligible effects on the general picture.

described by a simple model of exponential decline. When fitting such a model to this data, we get a predicted half-life of about 100 years for the stock of nobility.[24]

There are currently almost 600 remaining noble families. If the stock continues to decline at this rate, there will remain about 300 families by the year 2100, 150 families by the year 2200, and so forth. In sum, nobility as an institution appears to have considerable longevity, despite no longer enjoying economic or political privileges. If this simple model above comes close to predicting future developments, we ought to conclude that simply ceasing to create new noble families is not a very effective way of terminating the institution of nobility.[25]

Hereditary nobility is an anachronism. Our analysis can be extended, however, to other hereditary status markers which may perpetuate in modern times. For example, several Ivy League universities have adopted a so called "legacy policy", whereby the probability of acceptance is higher, *ceteris paribus*, if a parent has attended the same university (Karabel, 2005). To the extent that an Ivy League education is a positional good, alumni status might come to play a similar role in marriage markets in the future. We encourage future research in this area.

In sum, status affects economic outcomes, and is positively assorted with wealth in the marriage market. Given the hereditary nature of nobility, this mating pattern has consequences for the transmission of inequality. Not only does wealth beget wealth in the marriage market, but an indicator of past wealth does so, too. We end our discussion by noting that the marriage of wealth and status is an old and familiar theme. Well known representations in popular culture include Hogarth's *Marriage à la Mode*, a series of 19th century engravings featuring a cash poor aristocrat, Lord Squanderfield, who marries off his son to the daughter of a merchant. In di Lampedusa's (1966) novel *The Leopard*, Don Fabrizio, an ageing Sicilian prince, reluctantly marries off his orphaned nephew to the daughter of a local businessman with new-found wealth

[24] Let S(t) denote the stock of noble families at time t, normalized by population size. From a regression of log(S) on the time interval since 1900, we get a time coefficient of -0.0079 (OLS with robust standard errors, p-value < 0.001, R2 = 0.986). Entering this estimate into the model of exponential decline we get an expression for the evolution of the stock of nobility over time:

$$S(t) = S0e - 0.0079(t - 1900) \tag{5.1}$$

[25] This reasoning assumes no frequency-dependent advantage. If, for example, the marriage market valuation of nobility is an increasing function of the scarcity of nobility, a frequency-dependent nobility premium might arise. This would imply far greater longevity for nobility than is predicted by our exponential model. Exploring the ramifications of this more fully would constitute an exciting avenue for future research. It should also be noted that since no new noble lineages are created, the institution is likely to eventually cease to exist, although this might take a very long time. In stochastic models of population dynamics, lineages are typically either eliminated or go to fixation (so that the entire population stems from the same lineage). In the case of nobility, both outcomes would make the institution redundant.

but no pedigree. In a revealing episode, Don Fabrizio is told that the prospective bride's grandfather Peppe was known by an unflattering nickname: "Peppe 'Mmerda". The prince is taken aback by this information, but remains determined to press ahead with matters. *Non olet*, he reminds himself, *non olet*.[26]

[26] (Latin) *Pecunia non olet*: money does not smell.

6. References

Akerlof, G. A. and Kranton, R. E. (2000), "Economics and Identity", *Quarterly Journal of Economics* 115, 715-753.

Almenberg, J. and Dreber, A. (2009), "Assortative Mating in Age, Education, Income and Wealth. Sweden 1985-2004", Mimeo, Stockholm School of Economics.

Ball, S., Eckel, C., Grossman, P. J. and Zame, W. (2001), "Status in Markets", *Quarterly Journal of Economics* 116, 161-181.

Banerjee, A., Duflo, E., Ghatak, M. and Lafortune, J. (2008), "Marry for what? Caste and Mate Selection in Modern India", Mimeo, Massachusetts Institute of Technology.

Becker, G. S. (1973), "A Theory of Marriage: Part 1", *Journal of Political Economy* 81, 813-846.

Becker, G. S. (1991), *A Treatise on the Family*, Cambridge: Harvard University Press.

Becker, G. S., Kevin M. Murphy and Ivan Werning (2005), "The Equilibrium Distribution of Income and the Market for Status", *Journal of Political Economy* 113, 282-310.

Bisin, A. and Verdier, T. A. (2000), "Beyond the Melting Pot: Cultural Transmission, Marriage, and the Evolution of Ethnic and Religious Traits", *Quarterly Journal of Economics* 115, 955-988.

Bowles, S. and Gintis, H. (2002), "The Inheritance of Inequality", *Journal of Economic Perspectives* 16, 3-30.

Cannadine, D. (1990), *The Decline and Fall of the British Aristocracy*, New Haven: Yale University Press.

Clark, G. (2007), *A Farewell to Alms. A Brief Economic History of the World*, Princeton: Princeton University Press.

Drewianka, S. (2003), "Estimating Social Effects in Matching Markets: Externalities in Spousal Search", *Review of Economics and Statistics* 85, 409-423.

Elmslie, B. and Tebaldi, E. (2008), "So, What Did You Do Last Night? The Economics of Infidelity", *Kyklos* 61, 391-410.

Fahlbeck, P. E. (1898), *Sveriges Adel*, Lund: Gleerups.

Fernández, R., Guner, N. and Rogerson, R. (2005)"Love and Money: A Theoretical and Empirical Analysis of Household Sorting and Inequality", *Quarterly Journal of Economics* 120, 273-344.

Fernández, R. and Rogerson R, "(2001). Sorting and Long-Run Inequality", *Quarterly Journal of Economics* 116, 1305-1341.

Frank, R. H. (1985), "The Demand for Unobservable and Other Positional Goods", *American Economic Review* 75, 101-116.

Frank, R. H. (1999), *Luxury Fever*, New York: Free Press.

Frey, B. S. (2006), "Giving and Receiving Awards", *Psychological Science* 1, 377-388.

Hamermesh, D. S. and J. E. Biddle (1994), "Beauty and the Labor Market", *American Economic Review* 84, 1174-1194.

Huck, S. and Müller, W. (2000), "Why the Rich are Nastier Than the Poor – A Note on the Distribution of Wealth When Individuals Care for Payoff Differentials", *Kyklos* 53, 153-160.

Karabel, J. (2005), *The Chosen*, New York: Mariner Books.

Koopmans, T. C. and Beckmann, M. J. (1957), "Assignment Problems and the Location of Economics Activity", *Econometrica* 25, 53-76.

di Lampedusa, G. T. (1966), *The Leopard*. New York: Pantheon.

Mailath, G. J. and Postlewaite A. (2006), "Social Assets", *International Economic Review* 47, 1057-1091.

McGuire, M. T., Raleigh, M. J. and Brammer, G. (1982), "Sociopharmacology", *Annual Review of Pharmacological Toxicology* 22, 643-661.

Mealy, L. (1985), "The Relationship Between Social Status and Biological Success: A Case Study of the Mormon Religious Hierarchy", *Ethology and Sociobiology* 6, 249-257.

Moldovanu, B., Sela, A. and Shi, X. (2007), "Contests for Status", *Journal of Political Economy* 115, 338-363.

Raleigh, M. J. and McGuire, M. T. (1994), "Seratonin, Aggression, and Violence in Vervet Monkeys", in: R. Masters and M. T McGuire (eds.), *The Neurotransmitter Revolution: Seratonin, Social Behavior, and the Law*, Carbondale: Southern Illinois University Press.

Rundblad, L. (1999), "Adliga strategier" Master's thesis, Department of Economic History, Stockholm University.

Sabatini, F. (2008), "Social Capital and the Quality of Economic Development", *Kyklos* 61, 466-499.

Smith, A. (2000), *The Theory of Moral Sentiments*, Amherst: Prometheus.

Tao, H. (2008), "Attractive Physical Appearance vs. Good Academic Characteristics: Which Generates More Earnings?", *Kyklos* 61, 114-133.

Turke, P. W. and Betzig, L.L. (1985), "Those Who Can Do: Wealth, Status, and Reproductive Success on Ifaluk", *Ethology and Sociobiology* 6, 79-87.

Veblen, T. (1899), *The Theory of the Leisure Class: An Economic Study of Institutions*, New York: Macmillan.

CHAPTER 4

Prediction Markets in Science

with Thomas Pfeiffer and Ken Kittlitz[1]

Abstract. Prediction markets are powerful forecasting tools. They have the potential to aggregate private information, to generate and disseminate a consensus among the market participants, and to provide incentives for information acquisition. These capabilities can be of high value for the practice of scientific research. We report an experiment that examines the compatibility of prediction markets with the current practice of scientific publication. Our findings suggest that publication does not detract from the functionality of prediction markets. We argue that for integrating prediction markets into the practice of scientific research it might be of advantage to use subsidizing market makers and keep markets aligned with current publication practice. We discuss potential problems and solutions for making prediction markets operational in scientific research.

1. Introduction

A prediction market is a marketplace for contracts whose payoffs depend on the outcome of a future event. In a well-functioning market, contract prices can be interpreted as forecasts about the outcome of the event, derived from the beliefs of all participants in the marketplace. Contract types can be designed to elicit various aspects of the probability distribution associated with an event (see Wolfers and Zitzewitz, 2004, p. 109-110). One popular contract type pays \$1 if a specific outcome is realized, and \$0 otherwise. The price of such a contract can be interpreted as the predicted probability of that outcome occurring. In practice, prediction markets facilitate trading by generating standardized contract rules, and are typically organized so that the market forecast is salient and easily interpreted.

To the extent that market prices can be interpreted as collective forecast, prediction markets disseminate, or broadcast, information. The mapping from individual beliefs to market prices is potentially complicated because individuals may differ in their risk aversion, and in the availability of funds for betting (Manski, 2006, Wolfers and Zitzewitz, 2006). But in practice prediction markets have been found to generate

[1] We thank Consensus Point, Anna Dreber, and the CLER-LAB at Harvard Business School. Thomas Pfeiffer gratefully acknowledges support from Society in Science. Johan Almenberg gratefully acknowledges support from the Jan Wallander and Tom Hedelius Foundation and the Torsten and Ragnar Söderberg Foundations.

good predictions for events ranging from product sales and horse races to presidential elections (Chen and Plott, 2002; Arrow et al, 2008). By making information on collective forecasts available to a broader public, the dissemination property of prediction markets has the potential to generate social utility.

Prediction markets can also facilitate more complex information processing tasks. If different market participants have different, complementary pieces of private information, prediction markets have the potential to aggregate this information. Aggregation of dispersed information means that the market prediction is close to the forecast of a hypothetical trader in possession of all the information. A market that aggregates all available information is said to display strong efficiency (Muth, 1961; Fama, 1970). The information aggregation property is illustrated by an example from Plott (1988): An event has three mutually exclusive outcomes, X, Y, and Z. The price of a contract depends on which outcome is eventually realized. If half of the traders in the market are informed that the outcome will not be X, and the other half is informed that the outcome will not be Y, a market that is able to aggregate this information will forecast that the outcome of the event will be Z. This prediction differs from simply averaging the traders' initial beliefs. Information aggregation requires that traders learn from the market. Laboratory experiments suggest markets can accomplish information aggregation tasks reasonably well, though the details of the process are not fully understood (Plott and Sunder, 1988; Forsythe and Lundholm, 1990).

Because making reliable predictions is a key objective in science, prediction markets offer potential benefits to scientific research (Hanson, 1995, Hanson, 1999). Dissemination and aggregation properties of markets might be valuable because knowledge in scientific research is often highly decentralized. Different researchers and research groups typically have access to different pieces of information and have different background knowledge concerning a specific scientific question. Data can be ambiguous in their interpretation and are typically subject to statistical and systematic error that can be difficult to estimate (see, for example, Henrion and Fischhoff, 1986). When settling a research problem, this may lead to diverging opinions within the research community. Prediction markets yield a consensus estimate on a scientific question that is communicated to market participants as well as to outside parties such as funding agencies and policy makers. The information disseminated by a functioning market has the potential to be more precise than a consensus obtained from traditional methods such as performing a meta-analysis of more or less biased data from the literature, or sampling expert opinion. An internal prediction market on printer sales at Hewlett-Packard, for example, has been reported to give more precise forecasts than the sales department (Chen and Plott, 2002), and the film-related prediction market Hollywood

Stock Exchange has been found to generate more accurate Oscar winner predictions than both a poll of actual academy voters and a panel of experts (see Pennock et al, 2001). A reliable consensus within the research community is important in order to generate agreement on what the most important current open questions are and thus helps to allocate resources in an optimal way (Pfeiffer et al, 2009).

A functioning prediction market can be used to fund research. Scientists that invest in a well-designed research project will gain an information advantage, i.e., their beliefs will be more accurate than the current consensus. In a prediction market, this information advantage can be used to re-capture the investment. Incentives can be amplified by subsidizing market makers that turn the zero-sum payoff structure of typical markets into a nonzero-sum game. Institutions that fund scientific research might therefore use prediction markets as an alternative to established ways of allocating resources. Instead of distributing funds based on past performance of researchers, funding institutions can use subsidized markets to efficiently allocate funds to researchers according to their actual contribution to a research problem.

Incentives in scientific research are currently centered around the publication record of researchers. Current practice in publishing, however, is known to introduce biases in the body of published results (Rosenthal, 1979, Ioannidis, 2005) and often leads to delayed information disclosure. By rewarding an instantaneous, honest and unbiased disclosure of research findings, prediction markets can help to overcome problems with publication-based incentives. However, given the pivotal role of publishing in the current practice of science it is crucial for potential applications of prediction markets to ensure compatibility with publishing.

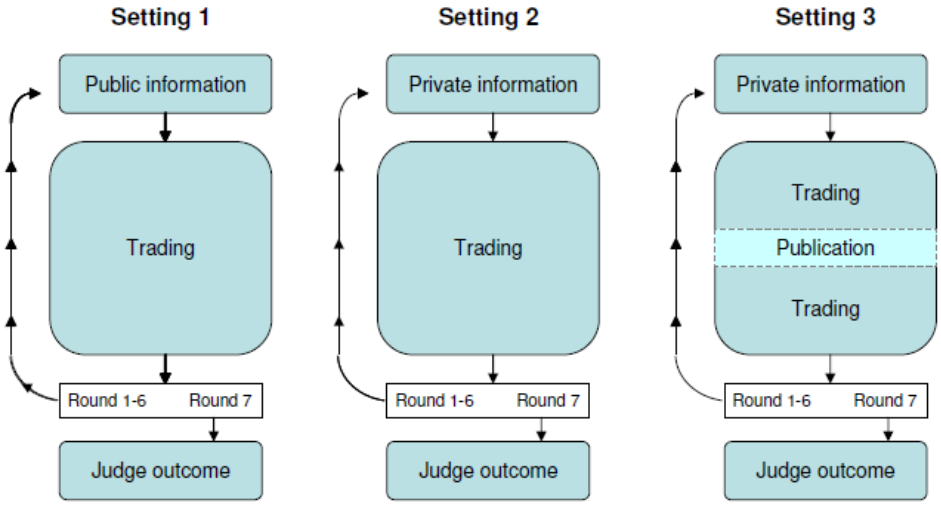
As a first step toward the integration of prediction markets into scientific research practices, we designed a simple experimental prediction market around a stylized process of scientific hypothesis testing and publishing. Hypothesis testing was presented to the participants within the context of a simple biochemical testing problem that has been used in previous experiments on decision-making in science (Pfeiffer et al, 2009). We presented six mutually exclusive hypotheses about a hypothetical biochemical pathway, and six different tests giving binary results that helped to identify the correct hypothesis. For each test, a positive result supported two of the hypotheses. The tests gave error-prone results, i.e., occasionally failed to provide a positive result for the true hypothesis (a *false negative*) or provided positive evidence for a false hypothesis (a *false positive*). The error rates were common knowledge. Further details on the hypotheses and tests are given in the Methods section.

In the experimental markets, participants traded contracts representing the six mutually exclusive hypotheses. Contracts were traded by 6 or 7 participants per market

using a web-based prediction market interface comparable to commercially available market platforms. A subsidizing market-maker based on a logarithmic scoring rule (Hanson, 2002; Hanson, 2003) was used to ensure liquidity despite the small number of traders, and to provide additional incentives for trading. Further details on the automated market maker are given in the Methods section.

Figure 1

Experimental Setup



The participants could trade during 7 consecutive trading rounds. Before each round, new information on the hypotheses was disclosed in the form of an error-prone test result. We investigated three different settings that differed in the way how information was distributed (see Fig 1). In Setting 1, information was always public. All participants received the same test result at the same time, and each disclosure was followed by a trading round in which they could buy and sell contracts. In Setting 2, information was always private. In each round, only one of the participants received a test result. This participant was chosen randomly before the experiment, and was determined independently for each round. Thus, some of the participants received more than one test result while others received no information over the course of an experiment. As before, each disclosure was followe by a trading round. In Setting 3, each piece of information was initially private, but eventually became public. One participant, drawn at random, received a test result before the trading round, just like in Setting 2. But halfway through the trading round the market was briefly suspended while the test result was published, i.e., disclosed to all participants. We used six different "information histories", each of which differed in the tests and results that were distributed. Each of these six information histories was used once in each setting, resulting in 18 different markets.

2. Methods

In total, 124 participants were recruited by the CLER-Lab at Harvard Business School. Most participants were students from the Boston area. Median age was 21, with about numbers of male and female participants. Participants received a performance-independent fee of $15 in addition to the payments earned in the experiment. The experiment was approved by Harvard University CUHS (F16041-101). Written informed consent was obtained from all participants.

The hypotheses used in the experiments were framed within the context of molecular biology. Our framing gives participants a concrete picture of what they are investigating, and avoids situations where they have prior expectations or preferences for the hypotheses under investigation. However, our findings are not specific to molecular biology and may be generalized across fields that engage in hypothesis testing. Suppose there are three genes, A, B, and C, that are known to interact in a linear biochemical pathway. The first gene activates the second, which in turn activates the third. The order of the sequence is unknown. Thus there are six possible pathways, ABC, ACB, BAC, BCA, CAB, and CBA, that form the set of hypotheses. Knowledge about these hypotheses can be characterized by six probabilities $p(h_1), \ldots, p(h_6)$.

For identifying the correct hypotheses, participants receive pieces of information from binary tests. The results indicate whether or not a specific gene activates another, i.e., whether A activates B, A activates C, etc. Thus there are six different tests (AB, AC, BA, BC, CA, and CB). Each test supports two of the hypothesis, and each hypothesis is supported by two tests. A positive result on test AB, for example, supports the sequences ABC and CAB, while sequence ABC is supported by positive results on test AB and BC. All of the tests are equally prone to type I and type II errors, and the error rates were common knowledge to all participants. We used the error rates $\alpha = 0.12$ and $\beta = 0.3$. These values are higher than the values of $\alpha < 0.05$ and $\beta < 0.2$ that researchers traditionally aim to achieve in the life sciences. We use these error probabilities to ensure that in the experiments, participants are exposed to errors at a considerable frequency. After a test has been disclosed, the probabilities associated with the hypotheses can be updated according to Bayes' Theorem.

With Bayesian updating, the posterior probabilities after test e_j are given by

$$p(h_i|e_j) = p(h_i)p(e_j|h_i) / \sum_i p(h_i)p(e_j|h_i) \tag{2.1}$$

for a positive test result, and

$$p(h_i|\tilde{}e_j) = p(h_i)(1 - p(e_j|h_i)) / \sum_i p(h_i)(1 - p(e_j|h_i)) \tag{2.2}$$

for a negative test result. The probability $p(e_j|h_i)$ of getting a positive result on test j given that hypothesis i is true equals $1-\beta$ if test j supports hypothesis i, and α if it does not support hypothesis i.

2.1. Trading platform and market maker. Participants used a web-based prediction market to trade contracts representing the six hypotheses. After an initial instruction period on a practice market, each participant received login details for a trading account that was funded with 100,000 virtual money units (henceforth VMU). This endowment was equivalent to USD 10. Contracts for the true hypotheses paid VMU 100 at the end of the experiment, whereas contracts representing any of the false hypotheses paid VMU 0. The trading platform used an automated marker maker. This is an algorithm that offers a buying price and a selling price at all times, thus ensuring that there is always a counterparty with which to trade. The market maker takes a risk, because the net portfolio of claims it buys and sells typically do not cancel each other out. We used a logarithmic scoring rule as the basis for the market maker. The algorithm uses the net sales $(S_1, \ldots, S_6)$ the market maker has done so far for each of the six claims to determines the prices for a infinitesimally small trade in claim i as

$$q_i = 100e^{S_i/b} / \sum_j e^{S_j/b} \tag{2.3}$$

where b is a liquidity parameter that also determines the maximum subsidy that may be provided by the market maker. In our experiment, we set $b = 2,000$.

2.2. Information histories. Information histories were generated by randomly choosing one of the 6 tests for each round and subsequently generating the test result based on the error probabilities given above. The resulting six histories are: History 1: BA false, CB false, CA false, BC true, AC false, AB true, CB true; History 2: BC true, BC true, CB false, AC false, BC false, CB false, CB False; History 3: AB true, AC true, AB true, AB true. CB false, BA false, CA false; History 4: CB false, AB true, CA false, BC true, BA false, BA false, CB false; History 5: BA false, CB false, BA true, AB true, AB true, BC false, BC true; History 6: BA false, BC true, BA false, CB true, CB false, AB true, BC true. To minimize the risk of contamination between settings, we changed the labels (i.e., the three letters) on the tests and hypotheses between each market and setting.

2.3. Informativity and support in the context of the settings. Let $\mathbf{P}$ be the vector of probabilities $(p_1, \ldots p_6)$ implied by the market prices. Suppose a trader whose beliefs are consistent with the current market $\mathbf{P}$ receives a new piece of information. He correctly updates his beliefs to $\mathbf{P}'$ and trades until market prices reflect his updated beliefs. If trading against a market maker like the one described in Hanson (2004) and

used in our experiment, the resulting profit is $b\log(p'_T/p_T)$, where p_T is the prior and p'_T is the posterior probability of the correct hypothesis. Therefore, in Setting 2, one would expect a traders' profit to be proportional to the support $S = \log(p'_T/p_T)$, at least in the absence of budget constraints. Of course, the trader does not know the correct hypothesis, and will therefore expect a payoff of $\sum_i p'_i \log(p'_i/p_i)$. To realize the profit, a trader has to wait until the market is judged, because unwinding the new positions by selling contracts merely moves prices back from $\mathbf{P}'$ to $\mathbf{P}$. In Setting 3, traders can in principle unwind their positions at a small loss once a piece of information has been made public, because in contrast to Setting 2 other traders should keep the prices close to $\mathbf{P}'$. Therefore, in this setting, traders could choose to realize a profit proportional to $\sum_i p'_i \log(p'_i/p_i)$ immediately after their information is disclosed, rather than waiting until the market is judged. The latter would yield an actual profit proportional to $S = \log(p'_T/p_T)$, equivalent to an expected payoff of $\sum_i p'_i \log(p'_i/p_i)$.

3. Results

As an example, one of the markets in Setting 1 is shown in Fig 2. We observe that participants trade at high frequency. On average, each participant traded more than 35 times over the course of the experiment, e.g., more than 5 times per trading round. Other markets showed similarly high trading activities. Participants traded even in the absence of private information, as observed in previous experimental asset markets (Smith, Suchanek and Williams, 1988; Copeland and Friedman, 1987). A lower trading volume than the one observed in the markets would have been sufficient to generate a correct pricing. Although in the market shown in Fig 2, all information is public and liquidity was high, we observe differences between the market prices and probabilities of the hypotheses as obtained by Bayesian updating. This mispricing likely reflects the participants' limitations in information processing and Bayesian updating. However, in general contract prices approximately followed rational pricing, and the final pricing was sufficiently precise for all participants to extract a net profit from the market maker.

To quantify mispricing and compare the overall performance of the market in the different settings we calculate the Kullback-Leibler divergence

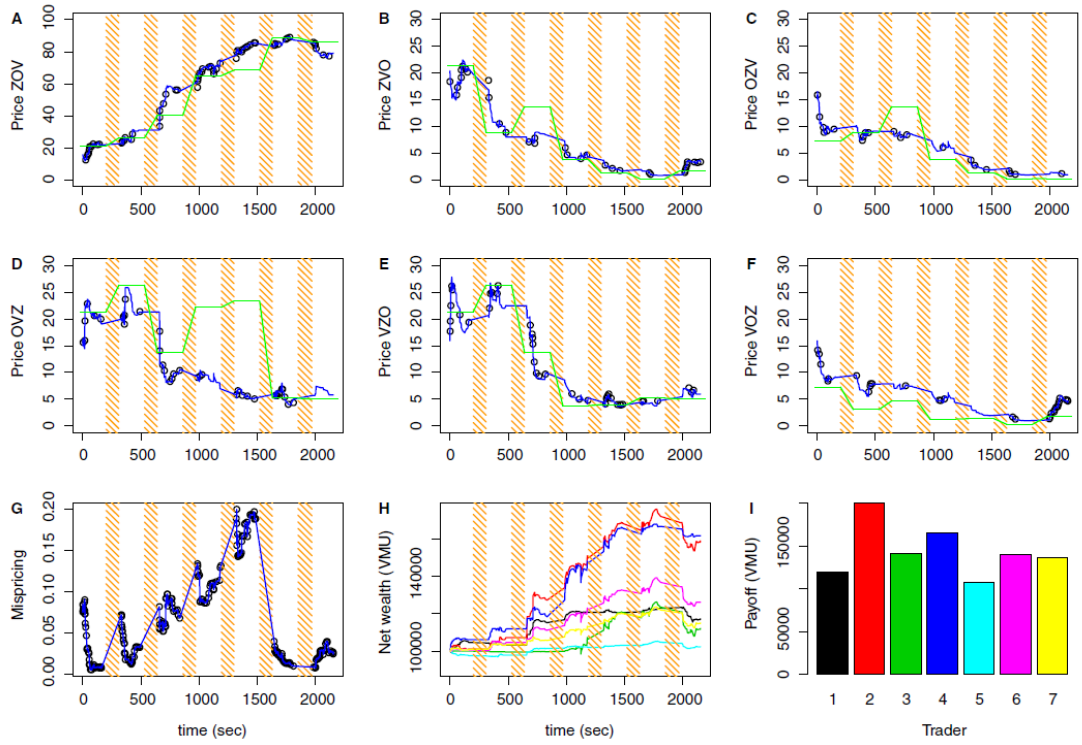
$$D_{KL}(\mathbf{P}||\mathbf{Q}) = \sum_i p_i log(p_i/q_i) \tag{3.1}$$

between a vector $\mathbf{P} = (p_1, \ldots, p_6)$ containing the probabilities of the hypotheses as obtained from Bayesian updating, and a vector $\mathbf{Q} = (q_1, \ldots, q_6)$ containing the actual prices of the contracts in the market using market maker prices (Kullback and Leibler, 1951). This measure of mispricing is proportional to the maximum arbitrage profit

that could accrue to a rational investor who knows the correct probabilities and trades against the market maker used in our experiments (Hanson, 2003, Chen, et al. 2007). Final mispricing, i.e. mispricing at the end of the last trading round, is shown for all 18 markets in Fig 3. To compare mispricing between different settings we use two-tailed paired t-tests on log-transformed final mispricing. We log-transform because a Shapiro-Francia test rejects normality of the mispricing but not of the log of mispricing (p-value 0.22). While performance in Setting 3 was similar to performance in Setting 1 (t-statistic: -0.15, p-value), mispricing was higher for Setting 2 (t-statistic: -3.8, p-value).

Figures 2A-I

Example Market



Figures 2A-I illustrate the workings of the market. Figures (A) – (F) show the prices of the six contracts over the course of one experiment session in Setting 1. Black dots indicate the trades in the contract. Black lines show the theoretical market maker price for infinitely small trades. The green line shows the actuarially correct pricing with Bayesian updating. The orange bars indicate that trading is suspended between rounds, while new information is disclosed. Figure (G) shows how mispricing evolved over time in this market, where mispricing is quantified as the Kullback-Leibler divergence between the probability distribution implied by the market prices and the

correct Bayesian posterior. Mispricing is typically generated when new information is released, and subsequently reduced by the traders (see for example trading round 1 and 2). As it becomes more and more likely that the first hypothesis (ZOV) is the correct one, the traders start overshooting in their estimate for the probability of this first hypothesis. This generates increasing mispricing in rounds 3-5. In round 6, the disclosed information reduces mispricing. During this round, reality (i.e. correct prices) catches up with the market forecast. Prices at the end of the last round give a good estimate of the correct probabilities. Figure (H) tracks the evolution of the net wealth of the seven traders participating in the market. After the last round, the hypotheses are judged. For each share of the correct hypotheses, participants receive a payoff of 100 VMU. This payoff is added to their cash position and determines the payoff. In this experiment all participants end up with a net gain, as shown in Figure (I). This illustrates the non-zero sum nature of trading due to the subsidizing market maker.

Within each setting, market efficiency did not depend on the overall trading volume, indicating that market efficiency was not inhibited by a lack of liquidity. A regression of mispricing on trading volume (both log-transformed) does not generate statistically significant results for any of three settings (p-value, trading volume: 0.38, 0.99 and 0.77 respectively).

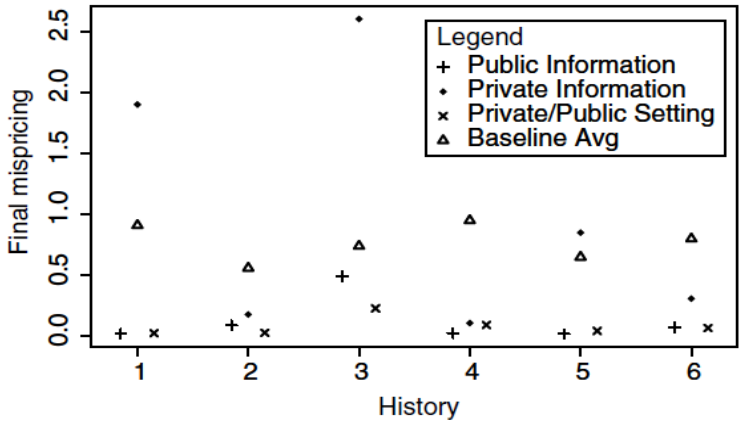
Because in Setting 1 and 3, all information becomes public at some point, the market disseminates public information, but does not need to aggregate private information. Thus in both settings final mispricing should only depend on the participants' ability to estimate the correct probabilities. One would therefore not expect systematic differences in the performance in both setting. This is indeed what we observe in the experiment. It did not make a difference for the quality of forecasts whether information was first private and then published, or was public from the beginning. Setting 2, in contrast, requires the aggregation of private information. This is much more complex task then the simple dissemination of information. In line with this, markets in Setting 2 show higher final mispricing for all of the six information histories.

In order to estimate whether the markets in Setting 2 aggregated private information, the observed market forecasts can be compared against the average belief of a set of rational individuals that all receive the same information as the market participants, do the correct Bayesian updating, but don't interact in a marketplace. In three of the markets, such an average belief would give a better forecast than market prices (see Fig 3). Thus, these markets clearly fail to aggregate information correctly. In the three remaining markets, prices produced forecasts that were considerably better than average belief. For these markets, performance falls into the range of the performance

in the settings with public information. This indicates that the markets are in principle capable of aggregating information, but that they sometimes fail to do so.

Figure 3

Mispricing after the final trading round



In Setting 2 and 3, participants have private information, at least for some time. For these settings we can investigate whether participants can gain a profit from having an information advantage. This is important if prediction markets are to provide incentives to invest into research to gain an information advantage. We assess the relation between the profit and the information received by a participant based on two metrics, namely the support that a test provides for the true hypothesis, and its informativity. Both quantities account for the fact that the value of a piece of information depends of the information that precedes it. To give an extreme example, if the true hypothesis is already known with certainty, then additional information is of no value.

The support that a test result gives for the correct hypothesis can be quantified as

$$S = log(p'_T/p_T) \tag{3.2}$$

where p_T is the prior and p'_T is the posterior probability of the correct hypothesis. Findings that maximize the posterior probability of the correct hypothesis are the most valuable ones. The support $S = log(p'_T/p_T)$ is proportional to the reward for increasing the probability of the correct hypothesis from p_T to p'_T, as provided by the market maker used in our experiment. However, this calculation requires knowing the correct hypotheses. By contrast, the expected value of S is given by the Kullback-Leibler distance between the prior and posterior probability distribution $\mathbf{P}$ and $\mathbf{P'}$,

$$D_{KL}(\mathbf{P}||\mathbf{P'}) = \sum_i p'_i log(p'_i/p_i) \tag{3.3}$$

This measure, which is the informativity of the information, is equivalent to the mispricing measure used above. Here, it quantifies mispricing from the viewpoint of a trader who received novel information, and determines the profit that can be gathered based on the information advantage. While $\log(p'_T/p_T)$ takes negative values for false findings, and positive findings for true findings, informativity is always non-negative, even if a test result is erroneous.

Calculating support for the correct hypothesis and informativity requires a prior probability. This prior could be taken from either correct Bayesian updating or from actual market prices. In a perfect market they are identical. In markets where mispricing is prevalent, it is suitable to use actual market prices because these prices likely provide a better representation of the actual beliefs of the traders. We therefore use market prices to calculate informativity I and support S of a test result.

Analysis of the payoff of the players reveals that the correlation between support S and net profit is stronger than the correlation between informativity and net profit. We find a positive correlation between support and payoff in Setting 2 (OLS, coefficient on support: 47,039, p-value: 0.04). Coefficient size should be related to the 100,000 virtual money units in each trading account at the outset of the experiment. For Setting 3, we observe a strong relation between support and profit (OLS, coefficient on support: 101,040, p-value: <0.01). This result means that in both settings, participants gained a profit from having an information advantage, but only if the information was correct.

In a post-experiment questionnaire, subjects self-reported their familiarity with Bayes' rule and in a separate question were asked to make a Bayesian inference. We generated a dummy variable that took on the value 1 if subjects reported being familiar with Bayes' rule and/or made the correct Bayesian inference. We observe a positive relation between this measure and net profit. The effect is statistically significant in setting 2 (coefficient: 56,867; p-value: 0.05), but not in setting 3. (coefficient: 3,620; p-value: 0.93). Subjects also reported their previous experience of betting on sports events or similar, as well as previous stock market experience. We combined this information to generate a dummy variable proxying for experience of gambling and/or trading. We observe a positive relation between this self-reported measure of experience net profit in the experiment, but the effect is not statistically significant (coefficients in Setting 1 and 2, respectively: 19,225; p-value: 0.40 and 32,271; p-value: 0.29).

Subjects could not self-select into market settings and were not assigned to settings according to their characteristics, so including subject characteristics should not affect our qualitative results. Indeed, our findings about the value of informativity, measured as the support for the correct hypothesis, are robust to the inclusion of the aforementioned variables. This also applies to the inclusion, in the regression analysis,

of age and gender variables. Allowing the average profits to differ between markets by including market fixed effects (i.e., a dummy variable for each market) improves the precision of our estimates but doesn't change our qualitative results. When regressing trader payoff on support, without control variables, the R^2 is about 0.11 for setting 2 and about 0.27 for setting 3. When including the aforementioned control variables this increases to about 0.56 and 0.30, respectively.

4. Discussion

There is a large literature on experimental asset markets in economics, dating back to the pioneering work of Chamberlin (1948) and Smith (1962). Previous studies have examined several aspects of our design in isolation, including the effect of private information (Forsythe, Palfrey and Plott, 1982; Plott and Sunder, 1982; Plott and Sunder, 1988; Forsythe and Lundholm, 1990), sequential information arrival (Copeland and Friedman, 1987) and Bayesian updating (Camerer, 1987; Smith, Suchanek and Williams, 1988; Peterson, 1993; Anderson and Sunder, 1995; Ganguly, Kagel and Moser, 2000; Hommes, Sonnemans, et al, 2008.). We extended this line of research by framing the information aggregation task as one of scientific discovery, and by using a computerized prediction market interface with an automated, subsidizing market maker. The framing of experiments can affect behavior (Liberman et al, 2004), as does the computerized interface (Williams, 1980). Our study is intended as a first step toward making prediction markets operational within the scientific community; for this reason we believe it is valuable to frame the task in these terms, and to use a technology which is as similar as possible to what would be used in the field. To the best of our knowledge there are no previous studies using the experimental economics toolkit to study the application of prediction markets to science.

In our experiments we tested three settings that differed in the way that information was disseminated. In Setting 1, the information was public at all times. Mispricing in this setting was low. Because no trader ever had an information advantage relative to the other traders, participants in this setting could only benefit from being faster than other traders at updating prices to reflect updated probabilities, following the publication of novel information, or from exploiting mispricing due to other traders' miscalculations. In Setting 2, different participants had different pieces of private information. The information was not made public at any point during the experiment. In this setting, mispricing in three of the six markets was comparable to the mispricing in the first setting. The other three markets, however, showed substantial mispricing at the end of the last trading round. Thus, information was not aggregated in an efficient, reliable manner in this setting. Nevertheless, participants could profit from

an information advantage. In Setting 3, participants received private information that was subsequently made public. In this setting, mispricing was as low as in Setting 1. In contrast to Setting 1, but similar to Setting 2, there was a clear positive relationship between net profits and having an information advantage.

Comparing the different settings allows us to draw some tentative conclusions about compatibility of prediction markets and publishing in the context of scientific research. The good performance of the markets in Setting 3 suggests that these markets combined the advantages of the two other Settings. Markets gave good forecasts about the probabilities of the hypotheses, while at the same time allowing participants to profit from information advantages. This indicates that combining publishing and prediction markets might be an attractive first step toward making prediction markets operational in science. Further theoretical investigations will be required to study under which conditions potential trading and publishing strategies are incentive-compatible.

Prediction markets on scientific issues are rare, despite the potential benefits. Aside from regulatory problems that have been discussed elsewhere (Bell, 2006, Arrow et al, 2008), this might be due to problems specific to the practice of science. In the following we discuss a number of problems and potential solutions.

Expertise, by definition, is held by a small number of people. Prediction markets, by contrast, function well when traders are numerous and make frequent trades. As our experiment shows, this problem can be mitigated through the use of automated market makers algorithms such as the logarithmic market scoring rule proposed by Hanson (2002). This market maker has the additional advantage of making it easy to subsidize the market, so that trading becomes a non-zero game. This provides an alternative way for funding institutions to distribute resources. At present there is not much experience of subsiding marker makers in practice. Further empirical research on market makers might help to establish prediction markets in scientific research.

Researchers cannot be expected to have extensive trading experience, and may not be used to thinking about prices as information signals. In laboratory experiments, the participants' prior trading experience has been found to be an important determinant of market efficiency (Forsythe, Palfrey and Plott, 1982; Plott and Sunder, 1982; Smith, Suchanek and Williams, 1988; Forsythe and Lundholm, 1990; Peterson, 1993; Anderson and Sunder, 1995). On the other hand, reasonably high numeracy levels can be expected among scientists, and a real-world prediction market offers much more time for reflection and planning that a laboratory experiment. And if participants in scientific prediction markets can be given some form of training, markets are likely to perform better. In addition, bets could be made by research groups, rather than individual researchers. This should reduce the effects of inexperience, albeit at the

cost of reducing the potential number of traders. Moreover, markets can be rational even if the participants, on average, are not: prices, and hence the market forecast, are driven by the marginal transaction and not by average beliefs (Camerer, 1987; Gode and Sunder, 1993; Jamal and Sunder, 1996; Ackert, Church and Shehata, 1997). Our experiment shows that if prediction markets are a complement to publication, markets remain functional even if information is not properly revealed by a researcher to the market.

Mispricing may arise due to erroneous interpretation of market signals as well as strategic attempts of participants to mislead other traders. Laboratory experiments have shown that when there is uncertainty about the prevalence of traders with inside information, traders may incorrectly infer that an uninformed trade is actually an informed trade, and move prices (and the associated forecast) inappropriately. Such bubbles, which have been dubbed "information mirages" (Camerer and Weigelt, 1991; Allen and Gale, 1992; Chakraborty and Yilmaz, 2008), are mistakes. Furthermore, theoretical work points to a separate question of the incentive compatibility of prediction markets. Chen et al. (2007) have shown that when the market maker is the logarithmic MSR, traders have incentives for truthful and immediate revelation of information if information signals are independent, conditional on the state of the world. This result does not extend to the case with conditional dependence, where bluffing and strategic reticence may become advantageous. However, the potential impact of both misleading errors and misleading strategic behavior is diminished when information is can be made public, suggesting that an added value arises from the combining prediction markets and conventional scientific publication. Participants may also attempt to manipulate the market in order to shed a favorable light on their own research. Experimental studies have shown that prediction market manipulation is difficult to achieve in practice (Camerer, 1998; Hanson, Oprea and Porter, 2006; Veiga and Vorsatz, 2009).

Finally, some open questions revolve around what type of scientific question that is suitable for trading in prediction markets. Scientific research evolves through the interplay of theoretical prediction and empirical testing. It would therefore be tempting to trade contracts contingent on the truth regarding a hypothesis. This is not practical, for several reasons. Prediction market contracts are typically bound to a specific future event. A hypothesis can usually not be judged with absolute certainty based on a single event. Moreover, when hypotheses are tested based on error-prone tests absolute certainty is never attained. Scientific research usually develops a consensus about a theory or hypothesis, rather than absolute certainty, and this consensus typically emerges over time rather than due to a specific event. One way around this problem is to use an agent that judges the hypothesis at a specific point of time. However, betting

on the correctness of a hypothesis, and betting on how the correctness of a hypothesis will be judged are two different things. A better solution might be to bet on the outcome of an experiment. This requires that the experiment is described in detail in advance – something that would be helpful for scientific research but at present is not common practice.

Contracts could also be based on the extent to which an experimental outcome agrees with a theoretical prediction. If a set of competing hypotheses yield different predictions on an outcome, criteria can be established that quantify how much an experimental outcome favors one hypothesis over the others. This simply requires a suitable judgment technology that is agreed upon in advance and is common knowledge to the market participants. Contract prices on the agreement of experiment and prediction might then be interpreted as a measure of how likely it is that a given hypotheses is true, relative to the likelihood of the other hypotheses. A good solution to settle problems with judging would be to bet on the results that will be published by a specific point in time. This is particularly suitable for fields where single experiments are inconclusive but standardized, so that overall estimates can be generated by pooling several experiments. Our experiments suggest that such linking of markets and publishing generates incentives for information advantages as well as reliable forecasts.

5. References

Ackert, L. F., Church, B. K. and Shehata, M. (1997), "Market behavior in the presence of costly, imperfect information: experimental evidence", *Journal of Economic Behavior and Organization* 33, 61-74.

Allen, F. and Gale, D. (1992), "Stock-price manipulation", *Review of Financial Studies* 5, 503-529.

Anderson, M. J. and Sunder, S. (1995), "Professional traders as intuitive Bayesians", *Organizational Behavior and Human Decision Processes* 64, 185-202.

Arrow, K., Forsythe, R., Gorham, R., Hahn, R., et al. (2008) "The promise of prediction markets", *Science* 320, 877-878.

Bell, T. W. (2006), "Prediction markets for promoting the progress of science and the useful arts", *George Mason Law Review* 14

Camerer, C. F. (1987), "Do biases in probability judgment matter in markets? Experimental evidence", *American Economic Review* 77, 981-997.

Camerer, C. F. (1998), "Can asset markets be manipulated? A field experiment with racetrack betting" *The Journal of Political Economy* 3, 457-482.

Camerer, C. F. and Weigelt, K (1991), "Information mirages in experimental asset markets", *Journal of Business* 64, 463-493.

Chakraborty, A. and Yilmaz, B. (2008), "Microstructure bluffing with nested information", *American Economic Review (Papers and Proceedings)* 90, 280-284.

Chamberlin, E. H. (1948), "An experimental imperfect market" *Journal of Political Economy* 56, 95-108.

Chen, K. and Plott, C. R. (2002), "Information aggregation mechanisms: concept, design and implementation for a sales forecasting problem" California Institute of Technology working paper.

Chen, Y., Reeves, D. M., Pennock, D. M., Hanson R. D., et al. (2007), "Bluffing and Strategic Reticence in Prediction Markets" *Proc. of the International Workshop on Internet and Network Economics (WINE)*, San Diego, CA.

Copeland, T. E. and Friedman, D. (1987), "The effect of sequential information arrival on asset prices: an experimental study", *Journal of Finance (Papers and Proceedings)* 42, 763-797.

Fama E. F. (1970), "Efficient capital markets: a review of theory and empirical work" *Journal of Finance (Papers and Proceedings)* 25, 383-417.

Forsythe, R. and Lundholm, R. (1990), "Information aggregation in an experimental market", *Econometrica* 58, 309-347.

Forsythe, R., Palfrey, T. R. and Plott, C. R. (1982), "Asset valuation in an experimental market", *Econometrica* 50:537-567.

Ganguly, A. R., Kagel, J.H. and Moser, D. V. (2000), "Do asset market prices reflect traders' judgment biases", *Journal of Risk and Uncertainty* 20, 219-245.

Gode, D. K. and Sunder, S. (1993), "Allocative efficiency of markets with zero intelligence traders", *Journal of Political Economy* 101, 119-137.

Hanson, R. D. (1995), "Could Gambling Save Science? Encouraging an Honest Consensus", *Social Epistemology* 9, 3-33.

Hanson, R. D. (1999), "Prediction markets", *IEEE Intelligent Systems* 14, 16-19.

Hanson, R. D. (2002), "Logarithmic market scoring rules for modular combinatorial information aggregation", George Mason University working paper.

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

Hanson, R. D. , Oprea R, Porter D (2006), "Information aggregation and manipulation in an experimental market", *Journal of Economic Behavior and Organization* 60, 449-459.

Hellwig M. F. (1980), "On the aggregation of information in competitive markets", *Journal of Economic Theory* 22, 477-498.

Henrion M. and Fischhoff, B. (1986), "Assessing uncertainty in physical constants", *American Journal of Physics* 54, 791–798.

Hommes, C., Sonnemans, J., Tuinstra, J. and van de Velden, H. (2008), "Expectations and bubbles in asset pricing experiments", *Journal of Economic Behavior and Organization* 67, 116-133.

Ioannidis, J. (2005), "Why most published research findings are false", *PLoS Med* 2(8):e124.

Jamal, K. and Sunder, S. (1996), "Bayesian equilibrium in double auctions populated by biased heuristic traders", *Journal of Economic Behavior and Organization* 31, 273-291.

Kullback, S. and Leibler, R. (1951), "On information and sufficiency", *Annals of Mathematical Statistics* 22, 79-86

Liberman, V., Samuels, S. M. and Ross, L. (2004), "The name of the game: predictive power of reputations versus situational labels in determining prisoner's dilemma game moves", *Personality and Social Pscyhology Bulletin* 30, 1175-1185.

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

Muth, J. F. (1961), "Rational expectations and the theory of price movements", *Econometrica* 29, 315-335.

Pennock, D. M., Lawrence, S., Giles, C. L. and Nielsen, F. A. (2001), "The power of play: efficiency and forecast accuracy in web market games", *Science* 291, 987-988.

Peterson, S.P (1993), "Forecasting dynamics and convergence to market fundamentals", *Journal of Economic Behavior and Organization* 22, 269-284.

Plott, C. .R and Sunder, S. (1982), "Efficiency of experimental security markets with insider information: an application of rational-expectations models", *Journal of Political Economy* 90, 663-698.

Plott, C. .R and Sunder, S. (1988), " expectations and the aggregation of diverse information in laboratory security markets", *Econometrica* 56, 1085-1118.

Pfeiffer, T., Rand, D. G. and Dreber, A. (2009), "Decision-Making in Research Tasks with Sequential Testing", *PLoS ONE* 4(2):e4607.

Rosenthal, R. (1979), "The file drawer problem and tolerance for null results", *Psychological Bulletin* 86, 638–641.

Smith, V. L. (1962), "An experimental study of competitive market behavior", *Journal of Political Economy* 70, 111-137.

Smith, V. L., Suchanek, G. L. and Williams, A. W. (1988), "Bubbles, crashes, and endogenous expectations in experimental spot asset markets", *Econometrica* 56, 1119-1151.

Veiga, H. and Vorsatz, M. (2009), "Price manipulation in an experimental asset market", *European Economic Review* 53, 327-342.

Williams, A. W. (1980), "Computerized double-auction markets: some initial experimental results", *Journal of Business* 53, 235-258.

Wolfers, J. and Zitzewitz, E. (2004), "Prediction markets" *Journal of Economic Perspectives* 18, 107-126.

Wolfers, J. and Zitzewitz, E. (2006), "Interpreting prediction market prices as probabilities", NBER Working Paper 12200.

CHAPTER 5

Do More Expensive Wines Taste Better? Evidence from a Large Sample of Blind Tastings

with Robin Goldstein, Anna Dreber, John W. Emerson, Alexis Herschkowitsch, and Jacob Katz[1]

Abstract. Individuals who are unaware of the price do not derive more enjoyment from more expensive wine. In a sample of more than 6,000 blind tastings, we find that the correlation between price and overall rating is small and negative, suggesting that individuals on average enjoy more expensive wines slightly *less*. For individuals with wine training, however, we find indications of a non-negative relationship between price and enjoyment. Our results are robust to the inclusion of individual fixed effects, and are not driven by outliers: when omitting the top and bottom deciles of the price distribution, our qualitative results are strengthened, and the statistical significance is improved further. These findings suggest that non-expert wine consumers should not anticipate greater enjoyment of the intrinsic qualities of a wine simply because it is expensive or is appreciated by experts.

1. Introduction

When symbolic content is an important part of consumption, the enjoyment of a good might become decoupled from its innate qualities. The symbolic content of a price tag has been emphasized in marketing research (e.g., Cialdini, 1998). At the same time, when goods with similar characteristics differ in price, a reasonable prior is that the more expensive good will, on average, be of a higher quality. People have been shown to expect a positive correlation between price and quality (e.g., Rao and Monroe, 1989). Consistent with this expectation, a meta-analysis reports positive correlations between price and quality ratings for most, but not all, of 1,200 product markets, but also finds that the range of these correlations is very large (Tellis and Wernerfelt, 1987).

[1] This paper originally appeared as the lead article in the *Journal of Wine Economics* 3(1), Spring 2008. Our findings were also discussed at length in Robin Goldstein's book *The Wine Trials* (Fearless Critic Media, 2008). We thank Jacopo Anselmi, Zoe Chance, Shane Frederick, Richard Friberg, Barry Goldstein, Erik Grönqvist, Daniel Horwitz, Roy Ip, Magnus Johannesson, Thomas Pfeiffer, Sue Stubbs and an anonymous referee for helpful comments and suggestions. Johan Almenberg thanks the Ragnar and Torsten Söderberg Foundations for financial support, and Johan Almenberg and Anna Dreber thank the Jan Wallander and Tom Hedelius Foundation for financial support.

For some goods, there is much heterogeneity in consumer tastes, making it harder to infer quality from revealed preferences. Nonetheless, a reasonable prior is that consumers on average will derive more enjoyment from the good with the higher price. Wine seems to be a good where consumer tastes are highly heterogeneous (Amerine and Roessler, 1976; Lecocq and Visser, 2006). While individuals may frequently disagree over which wine they prefer, the above hypothesis suggests a positive correlation between the enjoyment of a wine and its price.

A number of studies have reported positive correlations between price and subjective appreciation of a wine for wine experts (e.g., Oczkowski, 1994; Landon and Smith, 1997; Benjamin and Podolny, 1999; Schamel and Anderson, 2003; Lecocq and Visser, 2006). Non-experts, however, may not be particularly sensitive to some of the refinements that are held in high esteem by wine aficionados. Weil (2001, 2005) uses the following experimental setup: two bottles of wine are poured into four containers. Tasters are then given three of the containers and asked to distinguish which one differs from the other two. A random guess has 1/3 chance of being correct. In Weil (2001), the two wines are identical apart from year, but one wine is from a "good" vintage, and the other from a "bad" vintage.[2] The tasters get it right 41% of the time – only marginally better than a random guess.[3] In Weil (2005), the wines are a reserve bottling and a regular bottling, from the same producer and year. The fraction of correct answers is merely 40%. Moreover, Weil finds that even when tasters can distinguish between the vintages, they are about as likely to prefer the good one as the bad one. And among those who can distinguish the reserve bottling from the regular bottling, only half prefer the reserve. In both cases, the wines differ in price by an order of magnitude. These experiments highlight the discrepancy between experts and non-experts and the subjectivity of the wine experience.

Extrinsic factors, such as peer consumption and marketing actions, can also influence how a good is experienced. The price tag may in itself be such a factor. Recent research has shown that individuals appreciate the same wine more when they think that it is more expensive (Brochet, 2001; Plassmann et al., 2008). In other words, the price of a good affects the experienced utility derived from that good. Thus, to test the conjecture that the price of wine and the enjoyment of its intrinsic qualities are positively correlated, we need to examine the enjoyment of wine when individuals are unaware of the price.

[2] Weil uses pairs for which the famous wine critic Robert Parker has rated one of the bottles "average" to "appalling" and the other bottle "excellent" to "the finest".

[3] All of the significant difference is driven by the testers' ability to distinguish between the good and bad vintages from Bordeaux Pomerol.

In this paper, we use a large sample of more than 6,000 US blind tastings conducted by food and wine critic Robin Goldstein. We investigate the relationship between price and subjective appreciation of wines when the price is unknown to the taster. Subjective appreciation is measured by overall ratings assigned to wines by individual participants. Blind tastings offer the opportunity to isolate the experience of the wine itself from psychological confounds related to its price, presentation or published expert ratings.

Our main finding is that individuals who are unaware of the price do not, on average, derive more enjoyment from more expensive wine. In fact, unless they are experts, they enjoy more expensive wines slightly *less*. Our results are robust to the inclusion of individual fixed effects, and are not driven by outliers: when we omit the extremes of the price distribution our results are even more pronounced.

Our paper is organized as follows. In section 2, we describe our data. In section 3, we present our econometric model and report our results. We also perform a robustness check. We conclude in section 4, where we discuss some implications of our results and suggest directions for future research.

2. Data

The data set contains 6,175 observations from 17 blind tastings organized by Robin Goldstein. The blind tastings took place in the US between April 2007 and February 2008. In total, 506 participants tasted wine flights composed from 523 different wines. The wines were presented in a double-blind manner, so that neither the person serving the wine nor the person tasting the wine knew the identity, price, or any other characteristics of the wine aside from its color. Each taster assigned an overall rating to every wine tasted prior to discussing the wines with the rest of the group, and was not permitted to change his or her answer after discussion. The rating was the response to the question "Overall, how do you find the wine?" and the available answers were "Bad", "Okay", "Good", and "Great" In the data, these alternatives are coded from 1 to 4, with 1 corresponding to "Bad" and 4 corresponding to "Great".[4]

The price per bottle ranged from $1.65 to $150. The prices are average retail prices and were obtained from www.wine-searcher.com. The wines represent a broad variety of types (e.g., red, white, rosé, sparkling), country origins, and grapes.

The participants were unpaid volunteers from 21 to 88 years of age. Selection bias is a concern with any voluntary subject pool, and we have no reason to think that this

[4] Tasters ticked one of four boxes. In about 3% of the sample, tasters ticked in between two boxes, suggesting a rating somewhere in between the two responses. For simplicity, we dropped these observations from the regression. Including them makes no difference to our qualititative results, and the changes to the estimates are negligible.

is an exception. It is quite likely that the sample contains an over-representation of highly educated individuals, and an over-representation of individuals working in the food and wine industries. Nonetheless, the size of the sample and the general diversity of the tasters lead us to hope that inference will not be too restricted. For a more extensive description of the blind tastings, please see Chapters 8 and 9 and Appendix 1 in Goldstein (2008).

3. Results

Throughout the regression analysis, we use both an ordered probit estimator and a linear estimator (OLS). In both cases, we consistently use robust standard errors. The ordered probit estimator is particularly well suited to an ordinal dependent variable, but we find that OLS also performs well, and yields estimates that are easier to interpret. In any case, the two models generate highly consistent results. The dependent variable is the overall rating, measured on a scale from 1 to 4, with 4 being the highest rating. The main independent variable is the price variable, expressed as the natural logarithm of the average retail price per 750 mL of the wine in US dollars.[5]

In Model 1, we regress the overall rating assigned to wine i, by individual j, on the price of the wine. About 12% of participants had some wine training, such as a sommelier course. In Model 2, we allow for the possibility that these "experts" rate wines in a different manner. We include a dummy variable for being an expert, as well as an interaction term for price and the expert dummy. In a linear regression, this allows both the intercept and the slope coefficient to differ for experts and non-experts. In terms of the linear specification, we can write these two models as

$$y_i = \beta_0 + \beta_1 \ln(P_i) + \varepsilon_i \tag{3.1}$$

and

$$y_{ij} = \beta_0 + \beta_1 \ln(P_i) + \beta_2 EXPERT_j + \beta_3 \ln(P_i) \times EXPERT_j + \varepsilon_i \tag{3.2}$$

where P_i is the price of wine i, and $EXPERT_j$ is a dummy variable indicating if taster j has wine training. If individuals found that more expensive wines tasted better, the correlation between overall rating and price would be positive. In our sample, this is not the case: for both the ordered probit estimates and the OLS estimates, the coefficient on price is negative. In Model 1, the OLS coefficient is about -0.04, implying that a 100% increase in ln(price) is associated with a 0.04 reduction in the overall rating (Table 1, column (c)). The negative effect for more expensive wines is statistically significant.

[5] If we didn't do this, we would be expecting a one dollar increase to have the same effect at the $5 price level as at the $50 price level. We get similar qualitative results using the dollar prices, but the statistical significance of the coefficients is not as good (but still significant).

For non-experts, the relationship between price and overall rating is negative; for experts, however, this is not the case. Our estimates of Model 2 show that the correlation between price and overall rating is positive – or, at any rate, non-negative – for experts (Table 1, columns (b) and (d)). The price coefficient for non-experts is still negative, of about the same size as before, and with greatly improved statistical significance. The coefficient on the $\ln(P_i) \times EXPERT$ interaction term is highly statistically significant (ordered probit p-value: 0.017; OLS p-value: 0.015). For experts, the net coefficient on price is the sum of the two, i.e., about 0.11 for the ordered probit and 0.09 for OLS. This net coefficient is significantly different from zero, but only at the 10% level (ordered probit p-value: 0.099; OLS p-value: 0.095), despite the large sample size.

Table 1

Dependent Variable: Overall Rating

	Ordered probit		*OLS*	
	(a)	*(b)*	*(c)*	*(d)*
ln(Price)	-0.047 (0.039)**	-0.061 (0.013)**	-0.038 (0.038)**	-0.048 (0.012)**
ln(Price)*Expert		0.171 (0.017)**		0.138 (0.015)**
Expert		-0.558 (0.001)***		-0.448 (0.001)***
Constant			2.297 (0.000)***	2.337 (0.000)***
N	5986	5972	5986	5972
R^2/pseudo-R^2	0.000	0.002	0.001	0.005

Robust p-values in parentheses. * significant at 10%; ** significant at 5%; *** significant at 1%

In sum, we find a non-negative relationship between price and overall rating for experts. Due to the poor statistical significance of the price coefficient for experts, it remains an open question whether this coefficient is in fact positive.

How large are these price effects? The coefficients are of a moderate magnitude, but non-negligible, given that wine prices cover a large range – both in our sample and in the general wine market. Suppose we have two wines, A and B, and Wine A costs ten times more then Wine B in dollar terms. In terms of a 100-point scale such as that used by *Wine Spectator*, the OLS estimation of Model 2 predicts that non-experts will assign an overall rating that is four points lower for wine A, whereas experts will assign an overall rating that is seven points higher.[6]

In addition, the coefficient on the expert dummy is negative, quite sizeable, and statistically significant (OLS expert dummy coefficient: -0.448; p-value: 0.001). In

[6] If the dollar price increases by a factor of 10, ln(price) increases by about 2.3. Hence the predicted effect on the overall rating of tenfold increase in the dollar price is 2.3 times the ln(price) coefficient for non-experts and experts, respectively, adjusted to a 100 point scale.

other words, the OLS estimation of Model 2 consists of two linear relationships, one with a higher intercept (2.337) but a negative slope (-0.048), and one with a lower intercept (1.889) but a positive slope (0.090). The point where the two lines cross each other is the price level at which experts and non-experts are expected to assign the same rating. If we take the model literally, this point occurs at the price of \$25.70, i.e., ln(price) $\approx$ 3.25. At this price, the model predicts that both groups will assign a rating of about 2.2. Below this price, the model predicts that experts will assign lower ratings to a wine than non-experts, and vice versa.

We also test a third model, including individual fixed effects. In terms of the linear specification, Model 3 can be written as

$$y_i = \beta_0 + \delta_j + \beta_1 \ln(P_i) + \beta_3 \ln(P_i) \times EXPERT_j + \varepsilon_i \tag{3.3}$$

where δ_j is a dummy for each individual taster. Including individual fixed effects has very little effect on the qualitative results and the minor differences only serve to reinforce our earlier conclusions, as both the negative effect for non-experts and the positive effect for experts become slightly stronger. These results are presented in Table 2. For each of the four regressions in Table 2, a Wald test rejects that the fixed effects are jointly equal to zero by a wide margin (p-value < 0.001), suggesting that this is a suitable addition to the model.

Table 2

Individual Fixed Effects

	Ordered probit		*OLS*	
	(a)	*(b)*	*(c)*	*(d)*
ln(Price)	-0.070 (0.007)***	-0.089 (0.001)***	-0.050 (0.009)***	-0.064 (0.002)***
ln(Price)*Expert		0.209 (0.011)**		0.151 (0.013)**
Individual fixed effects	Yes	Yes	Yes	Yes
Constant			2.487 (0.000)***	2.183 (0.000)***
N	5986	5972	5986	5972
R^2/pseudo-R^2	0.080	0.081	0.181	0.182

Robust p-values in parentheses. * significant at 10%; ** significant at 5%; *** significant at 1%

To make sure that our results are not driven by wines at the extreme ends of the price distribution, we also run our regressions on a reduced sample, omitting the top and bottom deciles of the price distribution. Given the broad range of prices in the sample, this is an appropriate precaution. The remaining wines range in price from \$6 to \$15.

Using the reduced sample, we estimate Model 2 (Table 3, columns (a) and (c)) and Model 3 (Table 3, columns (b) and (d)). This produces consistent and even more pronounced estimates. The coefficient on price is still negative, and in each case larger than when using the full sample. The statistical significance of the coefficients improves further, and the R-squared is higher.

Table 3

Reduced Sample, With and Without Individual Fixed Effects[7]

	Ordered probit		*OLS*	
	(a)	*(b)*	*(c)*	*(d)*
ln(Price)	-0.225	-0.173	-0.182	-0.122
	(0.001)***	(0.019)**	(0.001)***	(0.025)**
ln(Price)*Expert	0.523	0.515	0.421	0.364
	(0.002)***	(0.006)***	(0.002)***	(0.009)***
Expert	-1.301		-1.044	
	(0.000)***		(0.000)***	
Individual fixed effects	No	Yes	No	Yes
Constant			2.622	1.910
			(0.000)***	(0.000)***
N	4817	4817	4817	4817
R^2/pseudo-R^2	0.003	0.094	0.007	0.206

Robust *p*-values in parentheses. * significant at 10%; ** significant at 5%; *** significant at 1%

In sum, we use the reduced sample to check the robustness of our results with regard to mid-range price levels. Based on the above, we conclude that our results are not only robust but in fact even more pronounced when omitting observations at the extremes of the price distribution.

4. Conclusion

The pleasure we get from consuming wine depends both on its intrinsic qualities such as taste and smell and external attributes such as price and presentation. One may argue that the former influences our subjective appreciation through a bottom-up process, where the sensory apparatus plays a key role, and that the latter works through a top-down process, where beliefs and expectations about quality are important determinants.[8]

[7] The reduced sample excludes the top and bottom deciles of the price distribution.

[8] This, in turn, might depend on ulterior motives such as status concerns. Wine as a status signal, and the prospect that expensive wine could function as a positional good, is discussed in Goldstein (2008), chapter 5.

In this paper we have explored the bottom-up effects by looking at how participants in blind tastings rate wines. We find that, unless they are experts, individuals who are unaware of the price enjoy more expensive wines slightly less.

There is a large relevant literature related to the influence of extrinsic signals on taste experience. Lee et al. (2006) look at how knowledge of a beer's ingredients (normal beer with added vinegar) can affect subjective appreciation. They show that the timing of the information plays a substantial role. One group of tasters is told about the vinegar, tastes the beer, and assigns ratings. A second group is told about the vinegar after tasting the beer, but before the ratings are assigned. On average, individuals in the first group assign significantly lower ratings, suggesting that informing participants about the vinegar influences the experience in itself. Using functional magnetic resonance imaging (fMRI), McClure et al. (2004) find that having the subject's favourite brand's name on a drink makes it taste better than if it is unlabeled. In another fMRI study, Plassmann et al. (2008) test whether marketing actions such as changes in the price can influence the experienced pleasantness of a product such as wine. Testers are given different wines that they are told differ in price. In reality, some of the wines are the same but simply presented with different prices. Prices are found to correlate positively with experienced pleasantness, measured through both subjective reports and fMRI scans.

Marketing provides one channel through which consumers can be influenced to buy certain wines. But it is not the only one: wine critics/experts may also play a role in affecting wine prices and shaping consumer preferences. For example, Hadj Ali et al. (2007) find a positive effect of wine critic Robert Parker's ratings on the price of Bordeaux wine.

There is, however, some research expressing scepticism towards wine ratings and their use for the average wine drinker. According to Quandt (2007), many wine ratings do not actually convey any information, nor is there substantial agreement in ratings by experts. Consistent with this view, Weil (2007) investigates whether wine descriptions by experts convey information to wine consumers. This is tested by having testers match wine descriptions to wines. In a similar setup to Weil (2001, 2005), tasters are asked to distinguish the odd one out of three different glasses of wine. Only about 50% of the participants in Weil (2007) can distinguish the odd one out, and of those who manage to do it, only about half can correctly match a wine critic's description of the wine with the wine itself – which is no better than a random guess.

Our results indicate another reason for why the average wine drinker may not benefit from expert wine ratings: he or she simply doesn't like the same types of wines as experts. This is consistent with Weil (2001, 2005), who finds that even among the

subset of tasters who can distinguish between good and bad vintages, or reserve or regular bottlings, they are as likely to prefer the "better" one as the "worse" one.

These findings raise an interesting question: is the difference between the ratings of experts and non-experts due to an acquired taste? Or is it due to an innate ability, which is correlated with self-selection into wine training?[9] Investigating this further would be a fruitful avenue for future research.

In sum, in a large sample of blind tastings, we find that the correlation between price and overall rating is small and negative. Unless they are experts, individuals on average enjoy more expensive wines slightly less. Our results suggest that non-expert wine consumers should not anticipate greater enjoyment of the intrinsic qualities of a wine simply because it is expensive or is appreciated by experts.

[9] For a further discussion, see Chapter 4 of Goldstein (2008).

5. References

Amerine, M. A. and Roessler, E. B. (1976), *Wines: Their Sensory Evaluation*, San Francisco: W.H. Freeman and Company.

Benjamin, B. A. and Podolny, J. M. (1999), "Status, Quality, and Social Order in the California Wine Industry", *Administrative Science Quarterly* 44, 563-589.

Brochet, F. (2001), "Chemical Object Representation In The Field of Consciousness", Working paper, General Oenology Laboratory, France.

Cialdini, R. B. (1998), *Influence: The Psychology of Persuasion*, New York: Collins.

Goldstein, R. (2008), *The Wine Trials*, Austin: Fearless Critic Media.

Hadj Ali, H., Lecocq S. and Visser, M. (2008), "The Impact of Gurus: Parker Grades and en Primeur Wine Prices", *Economic Journal* 118(529), 158-173.

Landon, S. and Smith, C. E. (1997), "The Use of Quality and Reputation Indicators by Consumers: The Case of Bordeaux Wine", *Journal of Consumer Policy* 20, 289-323.

Lecocq, S. and Visser, M. (2006), "What Determines Wine Prices: Objective vs. Sensory Characteristics", *Journal of Wine Economics* 1(1), 42-56.

Lee, L., Frederick, S. and Ariely, D. (2006), "Try It, You'll Like It − The Influence of Expectation, Consumption, and Revelation on Preferences for Beer", *Psychological Science* 17(12), 1054-1058.

McClure, S. M., Li, J., Tomlin, D., Cypert, K. S., Montague, L. M., Montague P. R. (2005), "Neural Correlates of Behavioral Preference for Culturally Familiar Drinks", *Neuron* 44, 379-387.

Oczkowski, E. (1994), "A Hedonic Price Function for Australian Premium Wine", *Australian Journal of Agricultural Economics* 38, 93-110.

Plassmann, H., O'Doherty, J., Shiv, B. and Rangel, A. (2008), "Marketing Actions Can Modulate Neural Representations of Experienced Pleasantness", *Proceedings of the National Academy of Sciences* 105(3), 1050-1054.

Quandt, R. E. (2007), "On Wine Bullshit: Some New Software", *Journal of Wine Economics* 2(2), 129-135.

Rao, A. R. and Monroe, K. B. (1989), "The Effect of Price, Brand Name, and Store Name on Buyers' Percetions of Product Quality: An Integrative Review", *Journal of Marketing Research* 36, 351-37.

Schamel, G. and Anderson, K. (2003), "Wine Quality and Varietal, Regional and Winery Reputations: Hedonic Prices for Australia and New Zealand", *Economic Record* 79(246), 357-369.

Tellis, G. J. and Wernerfelt, B. (1987), "Competitive Price and Quality Under Asymmetric Information", *Marketing Science* 6(3), 240-253.

Weil, R. L. (2001), "Parker v. Prial: The Death of the Vintage Chart", *Chance* 14(4), 27-31.

Weil, R. L. (2005), "Analysis of Reserve and Regular Bottlings: Why Pay for a Difference Only the Critics Claim to Notice?", *Chance* 18(3): 9-15.

Weil, R. L. (2007), "Debunking Critics' Wine Words: Can Economists Distinguish the Smell of Asphalt from the Taste of Cherries?", *Journal of Wine Economics* 2(2): 136-144.

CHAPTER 6

When Does the Price Affect the Taste? Results From a Wine Experiment

with Anna Dreber[1]

Abstract. We designed an experiment that examines how knowledge about the price of a good, and the time at which the information is received, affects how the good is experienced. The good in question was wine, and the price was either high or low. Our results suggest that hosts offering wine to guests can safely reveal the price: much is gained if the wine is expensive, and little is lost if it is cheap. Disclosing the high price before tasting the wine produces considerably higher ratings, although only from women. Disclosing the low price, by contrast, does not result in lower ratings. Our finding indicates that price not only serves to clear markets, it also serves as a marketing tool; it influences expectations that in turn shape a consumer's experience. In addition, our results suggest that men and women respond differently to attribute information.

1. Introduction

Much economic analysis assumes that price simply reflects market structure, but price can also be a marketing tool, for example if the price tag itself affects how a good is perceived (see, e.g., Cialdini, 1998). Textbook illustrations of supply and demand typically feature downward sloping demand curves. For most goods this is a highly plausible assumption. Price may have a positive effect on demand, however, when the good in question is used for the purpose of costly signaling. In the case of positional goods (Veblen goods), the purpose is to signal affluence and thereby assert high status (see, e.g., Frank, 1985, 1999; McAdams, 1992). A closely related example on the supply side is when increased monetary incentives crowd out intrinsic motivation for providing a service (Gneezy and Rustichini, 2000a, 2000b; Mellström and Johannesson, 2008). In this case the purpose may be to signal altruism and thereby achieve social esteem (Bénabou and Tirole, 2006; Ellingsen and Johannesson, 2008). In both cases the price

[1] We thank Emma von Essen, Magnus Johannesson and Eva Ranehill for great comments on the manuscript, and Thomas Pfeiffer, David G. Rand and Nils Wernerfelt for great assistance during the wine tastings. Johan Almenberg thanks the Jacob Wallenberg Foundation and the Torsten and Ragnar Söderberg Foundations for financial support, and Johan Almenberg and Anna Dreber thank the Jan Wallander and Tom Hedelius Foundation for financial support.

tag carries a semantic component, in the sense that it affects the signaling value of the commodity in question.

In practice, it can be hard to distinguish the signaling value of a high price from the tendency to associate high price with high quality. Consumers have been found to expect a positive correlation between price and quality (e.g. Rao and Monroe, 1989). Consistent with this expectation a meta-analysis has found positive correlations between price and quality ratings for most of the 1,200 product markets surveyed, but also that the range of these correlations is large, and even negative for some markets (Tellis and Wernerfelt, 1987). Consumers' perceptions of objective price-quality relationships have been found to be only moderately accurate (Lichtenstein and Burton, 1989), and the price-quality heuristic can be misleading, for example when goods of low quality are priced high (e.g., Cialdini, 1998).

In this paper, we address one particular good – wine – to shed some more light on the relationship between the price of a wine and the individual enjoyment of the wine. Specifically, we explore if, and how, information about the price of a wine affects the experience of tasting the wine.

Attribute information, such as the price or the ingredients of the good, has a more powerful effect on the perception of quality when the experience of the good is ambiguous (Hoch et Ha, 1986). Tasting wine is a relatively ambiguous experience for many consumers. Objective measures of wine quality are not easily defined, and consumer tastes with regard to wine are highly heterogeneous (Amerine and Roessler, 1976; Lecocq and Visser, 2006). Wine judges display low within-subject correlations when unknowingly judging the same wine multiple times (Hodgson, 2008).[2] Tasters are only marginally better than a random guess at distinguishing vintage years from non-vintage years from the same vineyard, or reserve bottlings from regular bottlings from the same vineyard and year, despite very large differences in price (Weil, 2001, 2005).[3] And in a large sample of blind tastings, Goldstein et al. (2008) find that more expensive wines fail to get higher ratings.

Previous research indicates that price information may be an important determinant of the experienced pleasantness of a wine (Brochet, 2001; Plassmann et al., 2008). Using functional magnetic resonance imaging (fMRI), Plassmann et al. (2008) conduct a within-subject study with 20 participants. Each subject tasted three wines multiple

[2] In fact, only about 10% of the judges were able to replicate their score within a single medal group. Moreover, when the judges were consistent this usually happened for wines that they disliked. This study is particularly interesting given that another study has found a positive relationship between price and medal status such that awards can influence a winery's economic success (Lima, 2006).

[3] In Weil (2005) subjects are to distinguish between a reserve bottling and a regular bottling, from the same producer and year. Among those who can distinguish between these two bottlings, only half prefer the reserve, whereas the wines differ in price by an order of magnitude.

times, but were not always told which wine they were tasting. Subjects believed they were tasting five different wines that differed greatly in price. Two of these wines were in fact duplicates of two of the other wines, but labeled with a different price tag. For the tasting observations where the subjects were unaware of the price, ratings did not differ between two samples of the same wine. By contrast, when the supposed price was disclosed, the price level was found to correlate positively with experienced pleasantness, measured through both subjective reports and fMRI scans.[4] This research highlights the potential role of marketing in shaping how we experience the goods that we consume.

Plassmann et al. (2008) do not ascertain whether expectations constitute the mechanism whereby price affects the tasting experience. We extend their analysis by using an alternative methodology from consumer research. Our aim is to shed more light on the price effect, and, in particular, to better understand the mechanism through which price information exerts influence on the tasting experience. Unlike in Plassmann et al. (2008), our setup relies on between-subject comparisons, and does not involve deception.[5]

Our approach combines an information treatment with a timing treatment. By varying both the provision and the type of extrinsic information, as well as the timing of this information relative to the first-hand experience of the wine, our experiment sheds light on how consumers use extrinsic information about the product in forming an opinion about it. A blind setting, in which the extrinsic information is not disclosed, is compared to a setting in which the information is disclosed before tasting, as well as a setting in which the information is disclosed after tasting.

A similar setup has been used in consumer research, applied to clothing, paper towels and ground beef (Hoch and Ha, 1986; Levin and Gaeth, 1988). It has recently been applied to beer by Lee et al. (2006) who look at how knowing about a "secret ingredient" (vinegar added by the experimenter) affects experienced pleasantness. All three studies find that extrinsic information provided prior to first-hand experience with the good in question has a significant effect of how the good is experienced, whereas extrinsic information provided after the experience does not.

We replace the beer in Lee et al. (2006) with wine, and replace information about a secret ingredient with information about the wine's retail price per bottle ($40 or $5).

[4] Subjects' brains were scanned while subjects tasted the wine. The results show that increasing the price of the wine increases blood-oxygen-level dependent activity in an area thought to encode for experienced pleasantness (the medial orbitofrontal cortex).

[5] It is not self-evident that labeling a $90 wine as a $10 wine captures the appropriate price effect, which would be the difference in subjective well-being experienced when tasting a wine without knowing the true price relative to tasting this wine when aware of the price.

Vinegar in beer is likely to be bad news about the beverage to the minds of most beer consumers. By contrast, our experiment was designed to allow for positive information (the wine is expensive) as well as negative information (the wine is cheap).[6] Thus, we focus on price as an attribute, an important element in marketing (e.g., Cialdini, 1998).

The first purpose of our study is to gage the magnitude of both the positive and the negative expectational effect. We hypothesize that individuals will assign a higher rating to the wine when they know its high price, relative to those tasting it without knowing the price. We assume that many consumers expect a $40 wine to be a highly pleasant experience. We hypothesize that individuals will assign a lower rating to the wine if they know the price and consider it to be cheap. We assume that many consumers will not expect a $5 wine to be a very positive experience.

Second, we expect the timing of the price information to make a difference. Hoch and Ha (1986), Levin and Gaeth (1988), and Lee et al. (2006) find that information has a significant effect only when disclosed prior to first-hand experience of the good in question. On the basis of this we expect the information about price to have a larger effect, relative to the blind condition, in the before condition than in the after condition. In other words, we expect individuals to give higher ratings to the expensive wine when they know about the high price before tasting, but not necessarily when finding out about the price after tasting, and similarly with the cheap wine we expect individuals to give lower ratings when they know about the price before tasting.

Third, we test whether there is a gender difference in how the price information matters. The possibility of a gender difference was not intended as the focus of our study. It is highly plausible, however, that concerns about identity and social image form part of a price effect. Gender differences in behavior are commonplace in the experimental economics literature in general (see Croson and Gneezy, 2009 for an overview) and a number of studies find that men and women respond differently to treatments designed to trigger social concerns (e.g., Griskevicius et al., 2007; Mellström and Johannesson, 2008; von Essen and Ranehill, 2009, Rand et al, 2009). Given this, we have no reason to expect the effect of price on experienced pleasantness to be the same for men and women. Plassmann et al. (2008) do not control for such gender effects, and we are not aware of any test of gender differences in how the timing of the (price) information affects experienced pleasantness.

We find that an expensive wine gets considerably higher ratings when tasters are informed about the high price before tasting, relative to tasting "blind" – but only from

[6] Whether the prices are perceived in this way depends partly on the subjects' spending habits. We address this issue later in the paper.

female tasters. By contrast, women that taste the wine before being told the price do not assign significantly higher ratings, suggesting that once they form a first-hand opinion the attribute information has little effect. For men there is no significant difference between any of the three conditions. A possible interpretation of this discrepancy is that men and women respond differently to attribute information, with men being less sensitive to such cues. Alternatively, this might point to differences in how men and women relate to wine, or status, or both.

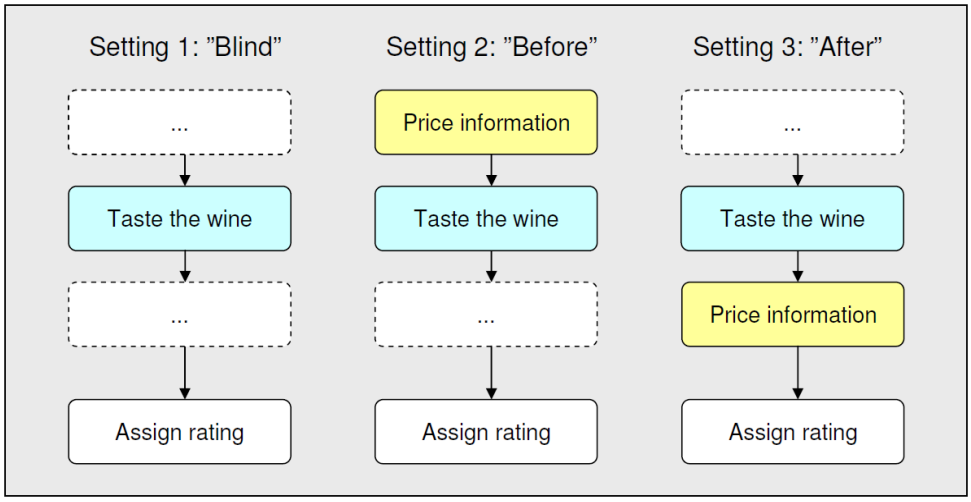
For the tasters that sampled the cheap wine, being informed about the price tag did not produce any noticeable changes in average ratings. This could point to an asymmetry between how positive and negative information shape perceptions of quality. A more likely explanation is that the bad news simply wasn't that bad: whereas the expensive wine was considerably more expensive than the tasters reported usually spending on wine, the cheap wine was simply in the lower range of typical spending. We elaborate on this later in the paper.

The outline of the paper is the following. We start by describing the setup of the experiment, then present our results, and finish the paper with a discussion.

2. Setup

All subjects followed the same procedure, illustrated in Figure 1 below. First, they received some information about the experiment. Next, they tasted one of two wines. The wine was either expensive or cheap. Finally, they received a short questionnaire, at the beginning of which they were asked to rate the wine. There were three information settings. In the "blind" setting, the price was not mentioned in the experiment. In the "before" setting, the price was mentioned in the information about the experiment, prior to tasting the wine. In the "after" setting, the price was mentioned at the top of the questionnaire, after having tasted the wine but still before rating it. Subjects were allocated randomly to one of the three information settings and one of the two wines. In other words, we use a between-subject design.

Figure 1
Experimental setup



Apart from the price, subjects received the same information in all three settings. They were told that the wine came from Portugal, that it was made out of a blend of different grapes, that they were to receive a glass of wine that they were to taste and that they subsequently would be asked to rate the wine.[7] In the actual tasting of the wine, subjects were given wine glasses filled with a small quantity of the wine and then given a few minutes to taste the wine. Once the subjects had indicated that they were done tasting, they were asked to set aside the glass until the experiment was over. Next, they were asked to assign a rating, using a visual analogue scale ranging from "undrinkable" to "perfection", with "OK" as the midpoint. Aside from this the scale was not labeled. Subjects were asked to circle a point (a tick mark) anywhere on the axis. In the statistical analysis we convert this to a 100 point scale.

3. Results

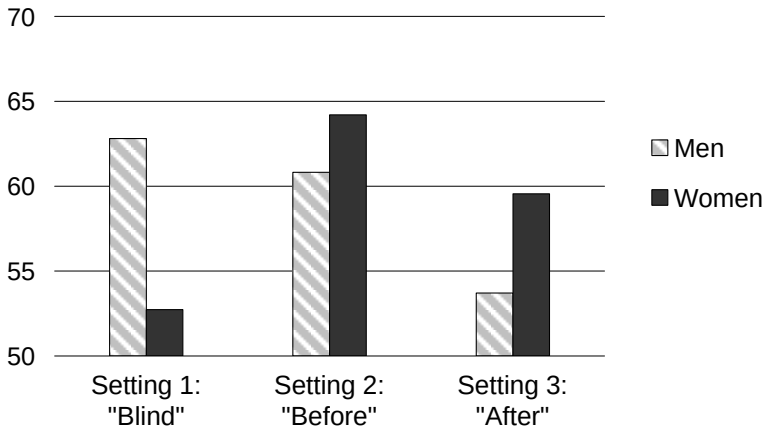
The study was conducted in Boston and Cambridge, Massachusetts, during the fall/winter semester of 2008-09. 135 individuals (40% women) tasted and rated a red wine with a retail price of $40, and 131 individuals (33% women) tasted and rated a red wine with a retail price of $5. The subjects consisted mainly of students and researchers at three universities. The average age was 29 (min: 21; max: 66).

[7] Subjects in the same session were randomized to different treatments. Making sure everybody read something made subjects not realize there were different treatments. In addition, we did not want subjects to sense that we were exploring the effect of the price tag. Embedding the price information among other information about the wine made this less obvious.

3.1. The expensive wine. Across all experimental settings and subjects the average rating of the expensive wine was about 59 out of 100. Average ratings by setting and gender are shown below.

Figure 2

Average rating of the $40 wine, by gender and experimental setting



The experimental data shows that the price can have a large effect on wine ratings, but this effect differs greatly between the sexes. Women, on average, assigned considerably higher ratings to the wine when they were informed about the $40 price tag before tasting. In terms of a 100 point scale, this effect implies that the rating increases by, on average, about 11.5 points. In terms of the visual analogue scale that subjects used for rating the wine, this effect represents about a quarter of the distance between "OK" and "perfection". The effect is statistically significant at the 5% level, regardless of whether we run the regression separately for both sexes or jointly, incorporating a dummy for being female as well as interaction terms for being female and the two information treatments. In the joint regression, the interaction term is statistically significant, and a Wald test rejects that the sum of the coefficients on "before" and the interaction term "before x female" is equal to zero (p=0.024). Men, by contrast did not assign higher ratings to the wine when they were informed about the price before tasting it.

Table 1
Experimental results for the expensive wine

	All subjects	Men	Women
Information about the price:			
Before *tasting (and rating)*	-2.00	-2.00	11.48
	(0.643)	(0.642)	(0.028)**
After *tasting (but still prior to rating)*	-9.11	-9.11	6.83
	(0.088)*	(0.088)*	(0.216)
Gender			
Female	-10.09		
	(0.039)**		
Gender interactions			
Before x *Female* [1]	13.47		
	(0.044)**		
After x *Female*	15.93		
	(0.037)**		
Constant	62.81	62.81	52.72
	(0.000)***	(0.000)***	(0.000)***
Observations	135	81	54
R^2	0.058	0.044	0.083

Robust p-values in parentheses. * significant at 10%; ** significant at 5%; *** significant at 1%

[1)] Note: a Wald test rejects that *Before* + *Before* x *Female* = 0. Prob > F = 0.024.

Neither women nor men assigned higher ratings to the wine when they were informed about the price tag after tasting. There is a noticeable tendency for men to assign lower ratings to the wine when they are told about the price after tasting. This effect is marginally statistically significant (p=0.09). Ten subjects, however, reported having some form of wine training, and if we extend our regression analysis to control for this the coefficient becomes smaller for men and seizes to be even marginally statistically significant (coefficient size: -7.40 instead of -9.10, p=0.185). We do not present this extended framework as our main model, because the number of subjects reporting wine training was small.[8] Moreover, they were all men. Nonetheless, this indicates that the negative effect for men in the after condition is not robust.

In other words, extrinsic information arriving after the subject has had first-hand experience of the good does not alter the subject's opinion of the good's quality. This is consistent with previous studies using the same design: Hoch and Ha (1986), Levin and Gaeth (1988), Lee et al. (2006) all find that information provided before experiencing the good has a significant effect on how the good is perceived, and that information provided afterwards does not.

[8] Controlling for expertise is justified, however, since it has previously been found that experts rate wines differently from non-experts (Goldstein et al., 2008).

3.2. The cheap wine. Across all experimental settings and subjects the average rating of the cheap wine was about 57. In the blind setting, the average rating was actually slightly higher for the 5$ wine than for the $40 wine (60.0 versus 58.5), in line with the finding in Goldstein et al. (2008) that most people do not prefer expensive wines, although this difference is not statistically significant.

For the cheap wine, we are unable to reject the null hypothesis that knowledge about the price has no effect on ratings, for either gender in any of the settings. Our data gives some indication of a corresponding negative effect of knowing about the low price of a cheap wine, but the absolute size of the effect is small and not statistically significant.

Table 2

Experimental results for the cheap wine

	All subjects	Men	Women
Information about the price:			
Before *tasting (and rating)*	-2.18	-4.74	1.15
	(0.66)	(0.37)	(0.87)
After *tasting (but still prior to rating)*	-7.13	-7.13	-3.08
	(0.19)	(0.19)	(0.72)
Gender			
Female	3.02		
	(0.67)		
Gender interactions			
Before x *Female*	-1.26		
	(0.86)		
After x *Female*	4.05		
	(0.69)		
Constant	58.98	58.98	62.00
	(0.000)***	(0.000)***	(0.000)***
Observations	131	87	43
R^2	0.031	0.023	0.008

Robust p-values in parentheses. * significant at 10%; ** significant at 5%; *** significant at 1%

In a post-experiment questionnaire, subjects reported spending on average $13 on a bottle of wine, with a standard deviation of about $6. Only two of 266 subjects reported spending $40 or more. Only about 5% reported spending more than $20. In the light of this, the $40 must be considered expensive relative to what the subjects usually spent on wine. By contrast, 16 subjects reported spending 5$ or less on average, and 40 % reported spending $10 or less.[9] Hence, the treatment effect of the low price

[9] There was no observable gender difference in spending behavior.

cannot be considered to be directly proportional to the treatment effect of the high price – i.e., it is possible that the cheap wine simply wasn't cheap enough. Subjects were asked to indicate their average weekly consumption of wine (number of glasses; frequency in parenthesis): < 1 (33%), 1-3 (40%), 4-6 (21%), 7-10 (6%), or > 10 (0%).

4. Discussion

Consumers' perceptions of objective price-quality relationships are not always very accurate, and this can have important implications. In the marketplace, consumers typically face vast amounts of information about the good they are about to consume. Price may be one of the more salient types of information, and if many people are not ready to expand time and effort to investigate the price-quality relationship, as suggested by Alpert (1971), then this leaves room for the use of price as an advertising tool, in a way that may be unrelated to the objective quality of the good.

We find that women assign considerably higher ratings to a wine if they are informed that it is expensive before tasting it. If they are informed that a wine is expensive after tasting it, assigned ratings are still higher than in the blind condition, but the difference is not statistically significant. When the wine is cheap, we do not find any negative effects of being informed about the price. For male tasters, we do not find any significant effects of knowing about the price – high or low – on average ratings.

Our main finding should surprise few: knowledge about the high price of a good can affect how it is experienced. In a world where luxury goods manufacturers routinely incorporate easily recognized logotypes into their designs, it can safely be assumed that knowledge about the high price of a good is considered a positive attribute that may confer status on its owner (e.g., Frank, 1999). In addition, many consumers use a price-quality heuristic that leads them to expect higher prices to be correlated with better quality. Tasting wine has been shown to be an ambiguous experience for many, if not all, consumers. Objective measures of wine quality are not easily defined; consumer tastes with regard to wine are highly heterogeneous. Extrinsic information, such as the price of the good, is likely to play a particularly important role when consumers are less confident in their own perceptions of quality.

In our view, the absence of a significant corresponding negative effect for a low price is most likely due to the design of our treatment, and not indicative of a deep asymmetry in how people react to high and low prices. In the post-experiment questionnaire, subjects reported their average level of spending on wine. The expensive wine was considerably higher than this average expenditure. The cheap wine, by contrast, was not below average expenditure in a way that can be considered proportional. In other words, most of our subjects typically consumed cheap wine. It is not surprising then

that being informed about the cheap price did not have significant effects on ratings. It could be argued that a $5 wine is probably more in line with what tasters in the blind setting are used to drinking and would expect to be offered, compared to a $40 wine.

With regard to the gender difference, our finding can be interpreted in two ways: (1) There is no gender difference. Either the female price effect is a false positive or the absence of a male price effect is a false negative. (2) Men and women respond differently to social cues, including status concerns regarding positional goods. It is not self-evident that men and women should have evolved to react the same way to such cues, and ample experimental evidence indicates that such differences exist. In our view, the second explanation is at least as plausible as the first, and merits further exploration.

Our study builds on previous research on the relationship between the price and the subjective experience of wine, in particular Goldstein et al. (2008) and Plassmann et al. (2008), through the application of a methodology used in marketing research. That marketing actions can affect the experience of a good is in itself not a novel finding. Marketing research has for a long time sought to schematize and empirically evaluate the interaction of top-down cognitive processes, to which extrinsic information is addressed, with bottom-up sensory processes, i.e., the experience of the intrinsic qualities of the good.

Attribute information may lead consumers to invest more effort when experiencing the good (Hoch and Ha, 1986). We did not control for the amount of time spent tasting the wine. It should also be noted that neither our study nor Plassmann et al. (2008) provides much detail about how expectations, once formed, interact with first-hand experience of a good. We do not know whether our subjects were actively searching for confirmatory evidence of an expensive/nice taste, or whether the wine simply tasted better during the actual tasting, such that the cognitive work on expectations occurred while processing the price information rather than while tasting. Future research should seek to shed more light on this process. We also encourage future research exploring whether our findings extend to other types of goods, and in particular whether the difference in how men and women respond to attribute information is product-specific and indicative of a more general difference in preferences between men and women.

5. References

Alpert, M. I. (1971), *Pricing Decisions*, Glenview, IL: Scott, Foresman and Company.

Amerine, M. A. and Roessler, E. B. (1976), *Wines: Their Sensory Evaluation*, San Francisco: W.H. Freeman and Company.

Bénabou, R. and Tirole, J. (2006), "Incentives and Prosocial Behavior", *American Economic Review* 96(5), 1652–1678

Brochet, F. (2001), "Chemical Object Representation in the Field of Consciousness," Working Paper, General Oenology Laboratory, Talence, France.

Cialdini, R. B. (1998), *Influence: The Psychology of Persuasion*, New York: Collins.

Croson, R. and Gneezy, U. (2009), "Gender Differences in Preferences", forthcoming in *Journal of Economic Literature*.

Ellingsen, T. and Johannesson, M. (2008), "Pride and Prejudice: The Human Side of Incentive Theory", *American Economic Review* 98(3), 990-1008.

von Essen, E. and Ranehill, E. (2009), "Dominance and Submission: Social status Biases Economic Sanctions", Working Paper, Stockholm School of Economics and Stockholm University.

Frank, R. H. (1985), "The Demand for Unobservable and Other Nonpositional Goods", *American Economic Review* 75(1), 101–116.

Frank, R. H. (1999), *Luxury Fever*, New York: Free Press.

Gneezy, U. and Rustichini, A. (2000a), "A Fine is a Price", *Journal of Legal Studies* 29(1), 1-18.

Gneezy, U. and Rustichini, A. (2000b), "Pay Enough or Don't Pay At All", *Quarterly Journal of Economics* 115(3), 791-810.

Goldstein, R., Almenberg, J., Dreber, A., Emerson, J. W., Herschkowitsch, A. and Katz, J. (2008), "Do More Expensive Wines Taste Better? Evidence from a Large Sample of Blind Tastings", *Journal of Wine Economics* 3(1), 1–9.

Griskevicius, V., Tybur, J. M., Sundie, J. M., Cialdini, R. B., Miller, G. F. and Kenrick, D. T. (2007), "Blatant Benevolence and Conspicuous Consumption: When Romantic

Motives Elicit Strategic Costly Signals", *Journal of Personality and Social Psychology* 93(1), 85–102.

Hoch, S. J. and Ha, Y. (1986), "Consumer Learning: Advertising and the Ambiguity of Product Experience", *Journal of Consumer Research* 13, 221–233.

Hodgson, R. T. (2008), "An Examination of Judge Reliability at a Major U.S. Wine Competition", *Journal of Wine Economics* 3(2), 105–113.

Lecocq, S. and Visser, M. (2006), "What Determines Wine Prices: Objective vs. Sensory Characteristics", *Journal of Wine Economics* 1(1), 42–56.

Lee, L., Frederick, S. and Ariely, D. (2006), "Try It, You'll Like It – The Influence of Expectation, Consumption, and Revelation on Preferences for Beer", *Psychological Science* 17(12), 1054–1058.

Levin, I. P. and Gaeth, G. J. (1988), "How Consumers are Affected by the Framing of Attribute Information Before and After Consuming the Product", *Journal of Consumer Research* 15(3), 374–378.

Lichtenstein, D. R. and Burton, S. (1989), "The Relationship between Perceived and Objective Price-Quality", *Journal of Marketing* 26(4), 429–443.

Lima, T. (2006), "Price and Quality in the California Wine Industry: An Empirical Investigation", *Journal of Wine Economics* 1(2), 176–190.

Mellström, C. and Johannesson, M. (2008), "Crowding Out in Blood Donation: Was Titmuss Right?", *Journal of the European Economic Association* 6(4), 845–863.

McAdams, R. H. (1992), "Relative Preferences", *Yale Law Journal* 102(1), 1–104.

Plassmann, H., O'Doherty, J., Shiv, B. and Rangel, A. (2008), "Marketing Actions Can Modulate Neural Representations of Experienced Pleasantness", *Proceedings of the National Academy of Sciences* 105(3), 1050–1054.

Rand, D G., Pfeiffer, T., Dreber, A., Sheketoff, R. W., Wernerfelt, N. C. and Benkler, Y. (2009), "Dynamic Remodeling of In-group Bias During the 2008 Presidential Election", *Proceedings of the National Academy of Sciences* 106(15), 6187-6191.

Rao, A. R. and Monroe, K. B. (1989), "The Effect of Price, Brand Name, and Store Name on Buyers' Perceptions of Product Quality: An Integrative Review", *Journal of Marketing Research* 36, 351–357.

Tellis, G. J. and Wernerfelt, B. (1987), "Competitive Price and Quality under Asymmetric Information", *Marketing Science* 6(3), 240–253.

Weil, R. L. (2001), "Parker vs. Prial: The Death of the Vintage Chart", *Chance* 14(4), 27–31.

Weil, R. L. (2005), "Analysis of Reserve and Regular Bottlings: Why Pay for a Difference Only the Critics Claim to Notice?", *Chance* 18(3), 9–15.

CHAPTER 7

Third Party Intervention, Group Size and Public Goods

with Anna Dreber, David G. Rand, and Coren L. Apicella[1]

Abstract. Costly third party punishment has been interpreted as a tool for studying the enforcement of social norms. Experiments on this topic typically involve a third party observer who can pay to decrease the payoff of one player who has behaved selfishly (or generously) toward another. We investigate whether third parties are sensitive to the number of players affected by this selfish or generous action. We allowed dictators to be selfish, fair, or generous, and unlike in other experiments, third parties could both punish and reward. Across all variations, responses followed a consistent and intuitive pattern: selfish behavior was punished while generous behavior was rewarded. Third party response was more pronounced when the dictator transfer had the nonrivalrous character of a public good, in the sense that both (a) the number of recipients increased and (b) the dictator's transfer was multiplied by a constant factor, so that the larger number of recipients did not reduce potential payoff of each recipient. Third party response did not change significantly when either of these manipulations was performed alone, suggesting a specific response to situations involving a public good.

1. Introduction

Shared beliefs of what constitutes appropriate behavior greatly affect human decision making in many social domains, ranging from dress codes and marriage practices to personal conflicts and public policy. Norms appear to be a human universal, with specific norms differing across cultures (Brown, 1991; Hauser, 2006). Depending on the norm, the same action can be classified as either good or bad (e.g., Ohtsuki and Iwasa 2006). Moreover, social norms may also contribute to the differentiation between groups and thus play a role in group selection (e.g., (Fehr and Fischbacher, 2004)).

It has been proposed that salient norms can be identified through examining the use of costly punishment by bystanders, or third parties, in experimental games (Fehr and Fischbacher, 2004). Third party punishment, or "moralistic punishment" (Kurzban, DeScioli et al., 2007), involves situations where I punish you after you defect against somebody else. In the type of non-repeated anonymous interactions that have been

[1] We are grateful for comments from Martin A. Nowak, Magnus Johannesson, Thomas Pfeiffer, David Cesarini, Drew Fudenberg, Hisashi Ohtsuki, Jorge M. Pacheco and Arne Traulsen. Johan Almenberg and Anna Dreber thank the Jan Wallander and Tom Hedelius Foundation, and Johan Almenberg thanks the Torsten and Ragnar Söderberg Foundations, for financial support.

explored, third parties can never benefit from choosing to reciprocate (i.e. punishing). Nonetheless, many people are willing to punish others at a cost to themselves in these scenarios (e.g., Fehr and Fischbacher, 2004).

Third party *rewarding*, a non-repeated encounter where I reward you because you have cooperated with somebody else, remains largely unexplored (a notable exception is Kahneman, Knetsch et al. 1986). Although similar in setup to an indirect reciprocity game, where my actions towards you depend on your previous actions towards others, here it is not possible to form a reputation. Hence it is never in one's self-interest to reward, and traditional indirect reciprocity theories predict that no rewarding will occur. A number of experimental studies have examined second party rewarding in settings where reputation formation is not allowed. The effects of rewards on cooperation and payoffs are mixed in both public good games (Walker and Halloran, 2004; Sutter, Haigner et al., 2006; Sefton, Schupp et al., 2007) and a two-player proposer game where the responder can either reward, punish, do both, or do neither (Andreoni, Harbaugh et al., 2003). In the latter game, the mean proposal is higher for rewards, whereas the modal offer is higher for punishment. These studies highlight a potential role for third party rewarding, something that we explore in this paper.

Third party punishment has been suggested as an effective tool for enforcing norms and promoting cooperation (e.g., Fehr and Fischbacher, 2004). It has been shown that third party punishment is more common in larger and more complex societies than in small-scale societies (Marlowe and Berbesque, 2008) Meanwhile, larger and more complex societies give rise to many situations where one individual can either extract a benefit at a cost to many others or incur a cost in order to benefit many others. Yet, to the best of our knowledge, how third party response to an action varies according to the number of individuals affected has not been investigated. That is the aim of this paper. More specifically, we are interested in how third parties respond to money transfers (or lack thereof) with a public good character. Importantly, we let third parties both reward and punish.

2. Methods

In our setup we use a modified dictator game. In the standard version of the dictator game, the dictator is an individual with an endowment to be allocated between herself and a recipient. The recipient does not have an endowment, and has no say over what allocation the dictator chooses. Numerous studies show that dictators frequently give a non-zero share to the other player (for an overview, see Camerer, 2003). The dictator game has been used to examine other-regarding preferences, such as altruism. As in Fehr and Fischbacher (2004), we introduce a third party that can affect the payoff of

the dictator, at a cost to herself. Across conditions, we let the dictator be selfish, fair or generous, and we allow third partiesto punish orreward.

In our version of the game, the dictator is at all times a single individual endowed with \$10. We vary the number of recipients, from 1 to 2 to 8 individuals. The dictator is presented with three options: give the entire \$10 endowment to the recipient(s), give half, or give nothing. Third parties are endowed with \$3 that can be spent to punish or reward the dictator, or can be kept by the third party. The response technology is 1:3, so that one dollar spent by the third party augments or reduces the dictator's payoff by three dollars.

We explore third party behavior in three settings. In Setting 1, the dictator's transfer is multiplied by n, the number of recipients, and then split equally between them. Each recipient thus gets the full dollar amount that the dictator chooses to transfer. In this setting, the money transferred by the dictator is like a public good for the recipients: each recipient experiences a benefit as a result of the dictator's action, and the size of this benefit is independent of the number of recipients.

Our experiment is designed to explore how strongly third parties respond to dictator actions in Setting 1. When the dictator gives something away, the multiplicative effect means that more money is "created" for the group as a whole, and the total group payoff is increased compared to when the dictator gives nothing. In other words, increasing the number of recipients introduces the possibility of "waste", which is not present in the standard, zero-sum dictator game. However, in principle waste and the number of recipients are two separate effects. In order to examine these effects in isolation from each other, we run the experiment in two additional settings.

In Setting 2, a single recipient receives a multiple z of the amount transferred by the dictator. We let z take on the values 1, 2, and 8. For each dollar kept by the dictator, the other player foregoes z dollars. For the same values of n, in Setting 1, and z, in Setting 2, the same amount is foregone by the group when the dictator keeps some or all of the endowment. Although increasing z creates a larger potential total group payoff, it also creates the potential for greater payoff inequity. When the dictator transfers money to the recipient, he creates additional wealth through the multiplier but at the same time creates an outcome where he is receiving a smaller payoff than the recipient. Previous work has shown that people care about both inequity and efficiency (Fehr and Schmidt, 1999; Engelmann and Strobel, 2004; Andreoni, 2007). This setting allows us to compare third party responses between treatments, holding the potential total group payoff constant.

In Setting 3, the number of recipients n takes on the values 1, 2, and 8, as in Setting 1. Here, however, the dictator transfer is not multiplied as in Setting 1, but simply

split equally between the recipients. Now, for one dollar transferred by the dictator, each of the n recipients gets $1/n$ dollars. This allows us to explore the possibility that third party responses are driven only by the number of individuals affected, regardless of the effect on total group payoff.

A total of 275 subjects from the Boston area participated voluntarily in one modified dictator game. The study was approved by the Harvard institutional review board. Written consent was obtained from all subjects before participating in the study. The subjects ranged in age from 18 to 70, were both students and non-students, and did not know the identity of any of the subjects that they were matched with. Subjects were not allowed to participate in more than one session of the experiment. All sessions were run in May 2007. Each subject was paid a show up fee of $5. All subjects were informed about the extensive form of the game, the endowments of each player, and that they would all receive a show-up fee. We used neutral language such as "add" or "subtract" money instead of "reward" or "punish". All subjects interacted anonymously, participated in a single treatment, and were only given one role (dictator, recipient, or third party) within that treatment. They were not informed about the other treatments. We elicited third party responses using the strategy method, as was done in previous work on third party punishment (Fehr and Fischbacher, 2004): Third parties were asked to specify their response to each option available to the dictator. Once the third parties had indicated their preferences, the dictator's actual choice was revealed. The advantage of this method is that it provides information about responses to outcomes that may occur very infrequently. Previous work has shown that using the strategy method may have a quantitative but not qualitative effect on decisions (Falk et al., 2005). We expect this effect on decisions to be orthogonal to the variation in conditions, allowing us to compare how third party behavior differs between treatments. In addition, the ratio of third parties to dictators was approximately 10:1. For each dictator, only one randomly selected third party's response was actually carried out. Third parties that were not matched with a dictator kept their $3 endowment, regardless of the choices they made. Third parties were fully informed of this design feature.

Throughout the data analysis we use a standard ordinary least squares estimator (OLS) with robust standard errors clustered by subject.

3. Results

Both third party rewarding and third party punishment occurred frequently in our experiment. As shown in Fig. 1, the majority of participants either rewarded, punished, or both. Of 184 total third parties, 66 subjects both rewarded and punished at least once (36%), 37 subjects rewarded at least once but never punished (20%), 17

subjects punished at least once but never rewarded (9%), and 64 subjects never did either (35%). Of all third party responses, 27% were rewards, 18% were punishments, and 55% were non-responses. Thus, not only did costly rewards occur, but they were more frequent than costly punishments.[2]

Figure 1

Third parties choices

Third party responses followed the same intuitive pattern in all treatments: as the amount transferred by the dictator increased, the average third party response also increased. A highly significant positive correlation between transfer and third party response exists in each treatment (Table 1). Selfish behavior (transferring nothing) was punished, and completely generous behavior (transferring everything) was rewarded. Transferring half the endowment was punished in some treatments and rewarded in others. However, transferring half was never punished more than transferring nothing, and never rewarded more than transferring everything.

Table 1

OLS regression of third party response on dictator transfer for each treatment.

Table 1.		*Setting 1*		*Setting 2*		*Setting 3*	
	Baseline	*n =2*	*n = 8*	*z =2*	*z = 8*	*n =2*	*n = 8*
Slope coefficient	0.13	0.20	0.25	0.20	0.18	0.16	0.12
P -value	0.004	<0.001	<0.001	<0.001	0.001	<0.001	0.004

Independent variable: dictator transfer. Dependent variable: third party response.

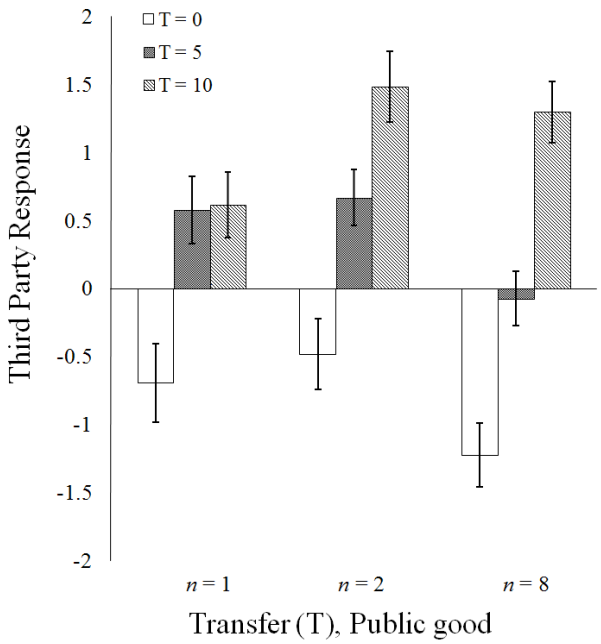
[2] Our study did not focus on dictator actions. We did not inform the dictators about the costly nature of third party actions, and thus cannot draw any conclusions from dictator behavior. However, 47% of the dictators chose to give away everything and 43% chose to keep half. Due to the small number of dictators (3 in each treatment), we cannot perform a substantial statistical analysis of differences between treatments.

Third party responses are increasing in the amount transferred by the dictator across all treatments, but this relationship may differ between treatments. To investigate this question, we perform an additional regression which includes a dummy variable and an interaction term for each treatment. Each treatment's dummy variable takes on the value 1 for data in that treatment, and 0 for all other data. This allows the regression intercept to differ between treatments. Each interaction term is the product of a treatment dummy and the explanatory variable (dictator transfer). It takes on the same value as the explanatory variable in one treatment and 0 in all other treatments. This allows the slope coefficient to differ between treatments. The p-values for the dummies and interaction terms give a direct measure of whether differences in the regression between treatments are significant. We are also able to distinguish between an overall level effect and a slope effect when comparing third party responses in the different treatments: a higher intercept indicates that third party responses are higher for all levels of dictator transfer, whereas a larger slope coefficient indicates that third parties responds more strongly to increases in the dictator transfer.

In Setting 1, each recipient got the full amount transferred by the dictator, regardless of the number of recipients. In this public goods setting, we found that third parties were sensitive to n, the number of individuals affected by the dictator's action (Fig. 2).

Figure 2

Third party responses in Setting 1 (public good)



Third parties rewarded generosity and punished selfishness significantly more when a dictator had 8 receivers than when a dictator only had 1 receiver (OLS regression with dummies and interactions; p-value 0.030, indicating a significantly steeper slope for the correlation between dictator transfer and third party response; see Table 2). However, the size of this effect is fairly moderate compared with the increase in possible group payoff and waste. Dictators transferring all \$10 were rewarded on average 2.1 times more when $n = 8$ as compared to $n = 1$. Dictators keeping all \$10 were punished on average 1.8 times more when $n = 8$ as compared to $n = 1$. In Setting 1, we also observed a change in the perception of a transfer of \$5, which resulted in the dictator and each recipient all receiving \$5. When $n = 1$ or $n = 2$, keeping half was on average rewarded (T-test, two-tailed; $n = 1$ response not equal to 0, p-value 0.0291; $n = 2$ response not equal to 0, p-value 0.0034). But when $n = 8$, this egalitarian allocation was no longer considered generous, and drew a neutral average response (T-test, two-tailed; $n = 8$ response not equal to 0, p-value 0.71). This suggests that a considerable fraction of third parties are sensitive to the number of individuals affected and potential group payoff. However, there is no statistically significant difference between $n = 1$ and $n = 2$ when considering responses to all three possible dictator transfers (OLS regression with dummies and interactions; p-value 0.253; see Table 2).

Table 2

OLS regression of third party response on dictator transfer for all treatments

Table 2.

Treatment		Slope	*P* -value	Intercept	*P* -value	Observations
Baseline setup						
n = 1	***z = 1***	**0.13**	**0.001**	**-0.49**	**0.065**	***78***
Treatment interactions and treatment dummies						
Setting 1						
n = 2	*z = 2*	0.07	0.253	0.06	0.868	*81*
n = 8	***z = 8***	**0.12**	**0.030**	**-0.77**	**0.027**	*81*
Setting 2						
n = 1	*z = 2*	0.07	0.24	-0.35	0.364	78
n = 1	*z = 8*	0.05	0.41	-0.01	0.987	78
Setting 3						
n = 2	*z = 1*	0.03	0.60	-0.14	0.676	81
n = 8	*z = 1*	-0.01	0.86	-0.02	0.953	78

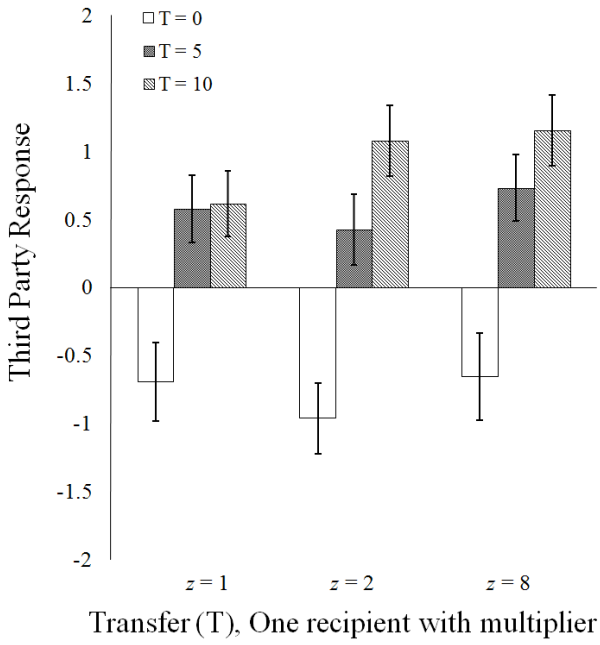
Dependent variable: third party response. Robust standard errors clustered on subject.

In Setting 2, a single recipient received a multiple z of the amount transferred by the dictator. For equal values of z in this setting, and n in the previous setting, the same total amount of money was foregone by the group for each dollar kept by

the dictator. This mimics the possibility of waste in Setting 1, without increasing the number of recipients (fixed at 1). If third parties were only concerned about the amount of money foregone by others as a result of the dictator's action, we should observe the same change in responses as in Setting 1. However, this was not the case (see Fig. 3). In fact, we found no significant change in third party responses as z increased (OLS regression with dummies and interactions; $z = 2$: p-value 0.240, $z = 8$: p-value 0.411; see Table 2). Yet in Setting 1, where the potential total group payoff was the same as it is here for equal values of n and z, there was a significant change in response going to 8.

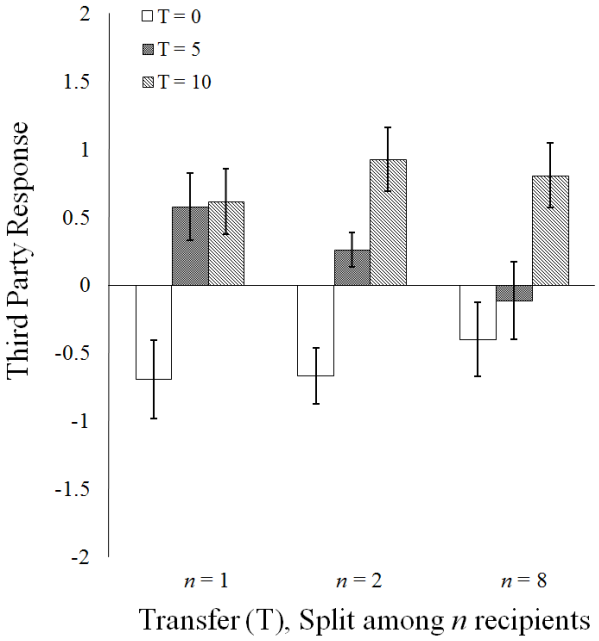
Figure 3

Third party responses in Setting 2 (single recipient)



The observed third party responses in Setting 2 allowed us to conclude that concerns about total group payoff alone do not seem to be driving third party responses. At this point, however, we have not ruled out the possibility that the results in Setting 1 were simply driven by a concern for the number of people affected, regardless of the extent to which they were affected. In our third and final setting, we show that this was not the case either.

Figure 4
Third party responses in Setting 3 (multiple recipients)



In Setting 3, we once again varied n, the number of individuals affected by the dictator's action. However, unlike in the first setting, here the transfer from the dictator was split equally by the n recipients. One more dollar for the dictator now meant $1/n$ dollars less for each recipient. We find no significant change across n values in third party responses (Fig. 4). The overall difference in third party responses when we varied n in this setting was not statistically significant (OLS regression with dummies and interactions; $n = 2$: p-value 0.60, $n = 8$: p-value 0.86; see Table 2).

4. Discussion

Many actions, especially in larger and more complex societies, can potentially influence the welfare of more than one other person. In economics, a benefit that is nonrivalrous (and non-excludable) is labeled a public good. Here we examine how third party response varies with the number of individuals affected by a dictator's action when the transfer has the nonrivalrous character of a public good. This gives us insight into how social norms, and their enforcement, may contribute to the provision and regulation of such goods.

We find that not only do third parties use costly punishment when given the opportunity, as has been found previously, but third parties also use costly rewarding. In fact costly rewards are more frequent than costly punishments. This suggests that

people are inclined to reward behavior that they deem good or generous at a cost to themselves. The preference for rewarding over punishing might in part be explained by a fear of retaliation. In most real life scenarios, both in early human societies (Marlowe and Berbesque, 2008) and today, punishers are not anonymous and so bear the risk of retaliation. Retaliation has been found to be common among animals (Clutton-Brock and Parker, 1995) and among humans in studies of cooperation games that allow for counter-punishment (Denant-Boemont, Masclet et al., 2007; Dreber, Rand et al., 2008; Nikiforakis, 2008). Thus, if individuals have limited resources to put toward enforcing cooperation it seems plausible that they may be better served by rewarding rather than punishing. In line with this, Kiyonari and Barclay (2008) find that in a public goods game followed by two rounds of targeted interactions, individuals who reward end up better off than those who punish.

When it comes to the number of individuals affected by the dictator in our experiment, previous work has produced related results, but in somewhat different contexts. Isaac, Walker and Williams (1994) compare groups of size 4, 10, 40 and 100 and show that group size under some circumstances correlates positively with the group's ability to provide the optimal level of a public good. Their results also show that cooperative behavior is influenced by a subtle interaction between group size and the marginal return of an individual's contribution that cannot be explained by either of these two things alone. The marginal return resembles our z, the multiplier of the dictator transfer. Another experiment on altruism and group size (Andreoni, 2007) finds that altruism in a modified dictator game is partly a *congestible* good, in the sense that people care both about the total benefit and the average benefit resulting from a monetary gift, with slightly more emphasis put on the former.

Our results, in combination with these previous findings, lead us to believe that the regulation of public goods plays a role in the evolution of cooperation that is, so to speak, greater than the sum of its parts. In our study, third party response is more accentuated when the number of recipients sharing a public good increases from 1 to 8. This increase contains two components: an increase in the number of individuals that are affected by the outcome, and an increase in the size of the aggregate transfer. In parallel treatments we show that third party response does not react to changes in one of these two components alone.

The experimental setup used in our study and in other third party studies is ultimately a non-repeated version of an indirect reciprocity game. The fundamental assumption of indirect reciprocity is that my potentially cooperative behavior toward you depends on what you have done to others. In such situations, it is often in the individual's self-interest to cooperate and reciprocate. Thus, it could be that the

propensity for costly intervention by third parties in non-repeated anonymous interactions is a misapplication of moral tendencies that evolved through indirect reciprocity. This possibility merits further study.

Whether third party norm enforcement has played an evolutionary relevant role in the provision of public goods remains an open question. As has been suggested previously, third party norm enforcement may create a group-level incentive for creating a public good (Fehr and Fischbacher, 2004). If this is the case, holding the action itself constant, helping or hurting more people should thus provoke more extreme third party responses. We find some evidence of this, but the effects are small. However, as Fehr and Fischbacher (2004) suggest, more than one third party may be needed to enforce a norm. In their view, this condition is probably met frequently in real life. Whether this conjecture holds true still remains to be explored.

5. References

Andreoni, J. (2007), "Giving Gifts to Groups: How Altruism Depends on the Number of Recipients", *Journal of Public Economics* 91, 1731-1749.

Andreoni, J., Harbaugh, W. T. and Vesterlund, L. (2003), "The Carrot or the Stick: Rewards, Punishments and Cooperation", *American Economic Review* 93(3), 893-902.

Brown, D. E. (1991), *Human Universals*, New York: McGraw-Hill.

Camerer, C. (2003), *Behavioral Game Theory*, Princeton, NJ: Princeton University Press.

Clutton-Brock, T. H. and Parker, G. A. (1995), "Punishment in animal societies", Nature 373(6511), 209-16.

Denant-Boemont, Masclet, L. D. and Noussair, C. N. (2007), "Punishment, counterpunishment and sanction enforcement in a social dilemma experiment", *Economic Theory* 33(1), 1432-1479.

Dreber, A., Rand, D. G, Fudenberg, D., Nowak, M. A. (2008), "Winners don't punish", *Nature* 452(7185), 348-51.

Engelmann, D. and M. Strobel (2004), "Inequality aversion, efficiency, and maximin preferences in simple distribution experiments", *American Economic Review* 94(4), 857-869.

Fehr, E. and Fischbacher, U. (2004), "Third-party punishment and social norms", *Evolution and Human Behavior* 25(2), 63-87.

Fehr, E. and Schmidt, K. (1999), "A theory of fairness, competition and cooperation", *Quarterly Journal of Economics* 114(3), 817-868.

Hauser, M. (2006). *Moral Minds*, New York, HarperCollins.

Isaac, R. M., Walker, J. M. and Williams, A. W. (1994), "Group size and the voluntary provision of public goods : Experimental evidence utilizing large groups", *Journal of Public Economics* 54(1), 1-36.

Kahneman, D., Knetsch, J. L. and Thaler, R. H. (1986), "Fairness and the Assumptions of Economics", *Journal of Business* 59(4), 285-300.

Kiyonari, T. and Barclay, P. (2008), "Free-riding may be thwarted by second-order rewards rather than punishments", *Journal of Personality and Social Psychology* in press.

Kurzban, R., DeScioli, P. and O'Brien, E. (2007), "Audience effects on moralistic punishment", *Evolution and Human Behavior* 28, 75-84.

Marlowe, F. W. and Berbesque, J. C. (2008), "More 'altruistic' punishment in larger societies", *Proceedings of the Royal Society B* 275(1634), 587-90.

Nikiforakis, N. (2008), "Punishment and Counter-punishment in Public Goods Games: Can we still govern ourselves?", *Journal of Public Economics* 92, 91-112.

Ohtsuki, H. and Iwasa, Y. (2006), "The leading eight: social norms that can maintain cooperation by indirect reciprocity", *Journal of Theoretical Biology* 239(4), 435-44.

Sefton, M., Schupp, R. and Walker, J. (2007), "The Effect of Rewards and Sanctions in Provision of Public Goods", *Economic Inquiry* 45(4), 671 - 690.

Sutter, M., Haigner, S. and Kocher, M. G. (2006), "Choosing the Stick or the Carrot? Endogenous Institutional Choice in Social Dilemma Situations", CEPR Discussion Paper No. 5497.

Walker, J. M. and Halloran, M. (2004), "Rewards and Sanctions and the Provision of Public Goods in One-Shot Settings", *Experimental Economics* 7, 235-247.

EFI, The Economic Research Institute

Published in the language indicated by the title.
A complete publication list can be found at *www.hhs.se/efi*
Books and dissertations can be ordered from EFI via e-mail: EFI.Publications@hhs.se

Reports since 2006

2009

Books

Engwall, Lars. *Mercury meets Minerva: business studies and higher education: the Swedish case.*
Henriksson, Lars. *Marknad eller reglering?: vägval för europeisk telekommunikation.* Forskning i Fickformat.

Dissertations

Hellström, Katerina. *Financial accounting quality in a European transition economy: the case of the Czech republic.*
Hernant, Mikael. *Profitability performance of supermarkets: the effects of scale of operation, local market conditions, and conduct on the economic performance of supermarkets.*
Lakomaa, Erik. *The economic psychology of the welfare state.*
Melén, Sara. *New insights on the internationalisation process of SMEs: a study of foreign market knowledge development.*
Törn, Fredrik. *Challenging consistency: effects of brand-incongruent communications.*
Wennberg, Karl. *Entrepreneurial Exit.*

2008

Books

Breman, Anna. *Forskning om filantropi. Varför skänker vi bort pengar?* Forskning i Fickformat.
Einarsson, Torbjörn. *Medlemskapet i den svenska idrottsrörelsen: En studie av medlemmar i fyra idrottsföreningar.* EFI Civil Society Reports.
Helgesson, Claes-Fredrik and Hans Winberg (eds). *Detta borde vårddebatten handla om.*
Jennergren, Peter, Johnny Lind, Walter Schuster and Kenth Skogsvik (eds). *Redovisning i fokus.* EFI:s Årsbok 2008. EFI/Studentlitteratur.
Kraus, Kalle. *Sven eller pengarna? Styrningsdilemman i äldrevården.* Forskning i Fickformat.
Petrelius Karlberg, Pernilla. *Vd under press: om medialiseringen av näringslivets ledare.* Forskning i Fickformat.
Portnoff, Linda. *Musikbranschens styrningsproblematik.* Forskning i Fickformat.
Sjöstrand, Sven-Erik. *Management: från kontorsteknik till lednings- och organisationsteori: utvecklingen på Handelshögskolan under 100 år: 1909–2009.*
Östman, Lars. *Den finansiella styrningens realiteter och fiktioner: de finansiella styrformernas svenska historia, berättelser om Petersson och "Ericsson", finansiell styrning – en ansats till generell teori.*
Östman, Lars. *Mycket hände på vägen från Buchhaltung till Accounting: delar av Handelshögskolan under 100 år.*

Dissertations

Axelson, Mattias. *Enabling knowledge communication between companies: the role of integration mechanisms in product development collaborations.*
Benson, Ilinca. *Organisering av övergångar på arbetsmarknaden: en studie av omställningsprogram.*
Elhouar, Mikael. *Essays on interest rate theory.*

Farooqi Lind, Raana. *On capital structure and debt placement in Swedish companies.*
Granström, Ola. *Aid, drugs, and informality: essays in empirical economics.*
Hvenmark, Johan. *Reconsidering membership: a study of individual members' formal affiliation with democratically governed federations.*
Höglin, Erik. *Inequality in the labor market: insurance, unions, and discrimination.*
Johansson, Marjana. *Engaging resources for cultural events: a performative view.*
Kallenberg, Kristian. *Business at risk. Four studies on operational risk anagement.*
Kviselius, Niklas Z. *Trust-building and communication in SME internationalization: a study of Swedish-Japanese business relations.*
Landberg, Anders. *New venture creation: resistance, coping and energy.*
Pemer, Frida. *Framgång eller fiasko? En studie av hur konsultprojekt värderas i klientorganisationer.*
Rosengren, Sara. *Facing clutter: on message competition in marketing communication.*
Schilling, Annika. *Kan konsulter fusionera?: en studie av betydelsen av identitet vid en fusion mellan konsultföretag.*
Schriber, Svante. *Ledning av synergirealisering i fusioner och förvärv.*
Sjödin, Henrik. *Tensions of extensions: adverse effects of brand extension within consumer relationship.*
Strandqvist, Kristoffer. *Kritiska år: formativa moment för den svenska flygplansindustrin 1944–1951.*
Strömqvist, Maria. *Hedge funds and international capital flow.*
Söderström, Johan. *Empirical studies in market efficiency.*
Sölvell, Ingela. *Formalization in high-technology ventures.*
Thorsell, Håkan. *The pricing of corporate bonds and determinants of financial structure.*
Ulbrich, Frank. *The adoption of IT-enabled management ideas: insights from shared services in government agencies.*
Östling, Robert. *Bounded rationality and endogenous preferences.*

2007

Books

Andersson, Per, Ulf Essler and Bertil Thorngren (eds). *Beyond mobility.* EFI Yearbook 2007. EFI/Studentlitteratur.
Einarsson, Torbjörn and Filip Wijkström. *Analysmodell för sektorsöverskridande statistik: fallet vård och omsorg.* EFI Civil Society Reports.
Ericsson, Daniel. *Musikmysteriet: organiserade stämningar och motstämningar.*
Samuelson, Lennart (ed). *Bönder och bolsjeviker: den ryska landsbygdens historia 1902–1939.*

Dissertations

Ahlersten, Krister. *Empirical asset pricing and investment strategies.*
Alexius, Susanna. *Regelmotståndarna: om konsten att undkomma regler.*
Andersson, Magnus. *Essays in empirical finance.*
Berg, Bengt Åke. *Volatility, integration and grain bank: studies in harvests, rye prices and institutional development of the parish magasins in Sweden in the 18th and 19th centuries.*
Bianchi, Milo. *Of speculators, migrants and entrepreneurs: essays on the economics of trying your fortune.*
Brodin, Karolina. *Consuming the commercial break: an ethnographic study of the potential audiences for television advertising.*
Elger, Max. *Three essays on investment-specific technical change.*
Hagberg, Axel. *Bankkrishantering: aktörer, marknad och stat.*
Hinnerich, Mia. *Derivatives pricing and term structure modeling.*
Hjalmarson, Hanna. *En växande marknad: studie av nöjdheten med konsumtionsrelaterade livsområden bland unga konsumenter.*
Hjelström, Tomas. *The closed-end investment company premium puzzle: model development and empirical tests on Swedish and British data.*

Kraus, Kalle. *Sven, inter-organisational relationships and control: a case study of domestic care of the elderly.*
Lindqvist, Erik. *Essays on privatization, identity, and political polarization.*
Macquet, Monica. *Partnerskap för hållbar utveckling: systrar av Oikos och guvernanten som blev diplomat.*
Melian, Catharina. *Progressive open source.*
Nilsson, Daniel. *Transactions in cyberspace: the continued use of Internet banking.*
Petrelius Karlberg, Pernilla. *Den medialiserade direktören.*
Portnoff, Linda. *Control, cultural production and consumption: theoretical perspectives, empirical dilemmas, and Swedish music industry practices.*
Sköld, Martin. *Synergirealisering: realisering av produktsynergier efter företagssammanslagningar.*
Sonnerby, Per. *Contract-theoretic analyses of consultants and trade unions.*
Tyrefors, Björn. *Institutions, policy and quasi-experimental evidence.*
Valiente, Pablo. *Re-innovating the existing: a study of wireless IS capabilities to support mobile workforces.*

2006

Books

Lundeberg, Mats, Pär Mårtensson and Magnus Mähring (eds) *IT & business performance: a dynamic relationship.* EFI Yearbook 2006. EFI / Studentlitteratur.
Thodenius, Björn. *Organisering av kunskap: en studie av Wallenberg Consortium North.* EFI Civil Society Reports.
Wijkström, Filip and Torbjörn Einarsson. *Från nationalstat till näringsliv?: det civila samhällets organisationsliv i förändring.*
Wijkström, Filip, Stefan Einarsson and Ola Larsson. *Staten och det civila samhället: idétraditioner och tankemodeller i den statliga bidragsgivningen till ideella organisationer.*
Östman, Lars. *Lysande ögonblick och finansiella kriser: Dramaten under ett sekel.*

Dissertations

Argenton, Cedric. *Quality provision in duopoly.*
Beckerman, Carina. *The clinical eye: constructiong and computerizing an anesthesia patient record.*
Borglund, Tommy. *Aktievärden i fokus: internationell påverkan på intressentrelationer vid förvärv och fusion.*
Breman, Anna. *The Economics of altruism, paternalism and self-control.*
Edquist, Harald. *Technological breakthroughs and productivity growth.*
Eklund, Jana. *Essays on forecasting and bayesian model averaging.*
Frostenson, Magnus. *Legitimitetskontrollen: en studie av etiska värderingars roll i gränsöverskridande förvärv och fusioner.*
Gaspar, Raquel M. *Credit risk and forward price models.*
Gustafsson, Peter. *Essays on trade and technological change.*
Hopkins, Elisabeth. *Is a higher degree of local currency pricing associated with lower exchange rate pass-through?: a study of import pricing in 51 Swedish industries.*
Kling, Ragnar. *Developing product development in times of brutal change.*
Langenskiöld, Sophie. *Peer influence on smoking: causation or correlation?*
Lychnell, Lars-Olof. *"Och fungerar det inte, gör vi på något annat sätt": en klinisk fallstudie av IT-relaterat förändringsarbete i småföretag.*
Meitz, Mika. *Five contributions to econometric theory and the econometrics of ultra-high-frequency data.*
Mendicino, Caterina. *Financial market imperfections, business cycle fluctuations and economic growth.*
Ovanfors, Anna. *Essays on nonlinear time series analysis and health economics.*
Paltseva, Elena. *Essays on commitment and inefficiency in political economy.*

Rogberg, Martin. *Den modeföljande organisationen: om acceptansen av TQM och andra populära managementmodeller.*

Silvennoinen, Annastiina. *Essays on autoregressive conditional heteroskedasticity.*

Sjögren, Ebba. *Reasonable drugs: making decisions with ambiguous knowledge.*

Slinko, Irina. *Essays in option pricing and interest rate models.*

Wilander, Fredrik. *Essays on exchange rates and prices.*