In this dissertation, four essays examine various ways in which people make decisions based upon others’ characteristics or actions. Individuals certainly can choose without regard to others and many decisions essentially lack social influences, such as the colour of one’s tie or scarf or car, but oftentimes peers may react to a choice in a way which decision-makers prefer not to disregard. Thus, some smoke or vote for the Greens because others do, or wear a certain haircut because it is the style at the time. Sometimes decisions may be private within some bound, such as choosing any uni-coloured car, but not a hippie-style multi-coloured one.

This dissertation seeks to demonstrate the influence of such social elements on choices pertaining to four distinct areas: (1) whether to seek public assistance (the dole); (2) whether to vote or abstain; (3) what to eat when others may observe one’s choice; and (4) whom to marry to avoid some causes of marital disharmony.

The treatment is always applied and mainly empirical; only the fourth essay makes use of an exclusively theoretical argument. This essay also differs in reducing the scope of social influences to characteristics of only one’s spouse(s). The article suggests an empirical pattern in which husbands tend to out-earn their wives in order to be able to withhold income from them to reduce the probability of non-paternity, with differences between the sexes declining as total incomes rise.

The other three articles proceed in the opposite way, finding empirical patterns and suggesting explanations for them. The first essay finds that welfare offices in buildings with features that enhance the visibility of entry (such as a set of steps before the entrance), tend to approve a greater share of applications than do other welfare offices. The suggested explanation is that only the neediest individuals will accept the risk of being seen to be ‘welfare cases’, eliminating comparatively haphazard applications to raise the approval rate.

The second essay looks at data from countries practicing multiple concurrent national referenda, finding a pattern in which more concurrent referenda are associated with lower turnout, akin to quantity discounts being associated with fewer sales. This puzzling relationship, too, may be accounted for by peer effects. If groups are affected by referenda outcomes but individuals do not really wish to vote, group pressure may be less effective when several groups care about different referenda, because each group wants to avoid pressuring the non-voter.

The article analyzing dietary decisions finds that customers at a restaurant in Western Finland consume lighter meals when their body types are bigger than those of their peers. This finding points to a mechanism whereby thinness is prized and people wish to signal conformity to thinness-inducing habits by foregoing a larger meal in exchange for social recognition.
Essays on Social Economics

Key words: Status and Peer Effects; Building Characteristics; Voting; Caloric Intake; Relative Income within Households

© Hanken School of Economics & Mats Ekman, 2018

Mats Ekman
Hanken School of Economics
Department of Economics
P.O.Box 479, 00101 Helsinki, Finland
PREFACE

I got a PhD! Only in Economics could I get a PhD. Slanderous statement at my fellow economists? Not at all! My degree is really only a testament to the breadth of our discipline, and perhaps to a greater willingness among its members to practise what they preach and focus on what few things lie within one’s capacity, rather than on the myriads of things one does poorly. For me, I believe this has been particularly beneficial. Other disciplines within social science would surely not welcome my fondness for rational choice, and outwith social science I could really only muster the interest to pursue a PhD in Philosophy or possibly in Linguistics, where professional success often comes with relatively unattractive financial rewards. So yes indeed, only in Economics could someone like me scrape through.

Although it is obvious that I would say it, I truly believe that my PhD is a felicitous match of interest and field. Of the fairly diverse interests that I have, the workings of society go as far back as I can remember, and at least as far as any of my other interests. From the implicit pricing of differently-sized marbles in the playground to the seemingly astounding fact that the car deck of a ferry we would take on family holidays was usually packed to capacity yet never left anyone behind on the quay, the sense that there must be some method to how societies work would fuel endless and fascinating – albeit not always fruitful – speculation.

As I became acquainted with Economics, what I think of as its three fundamental forces of society – Methodological Individualism, Markets, and Marginalism: the three M’s – revealed to me a method of analysis which continues unabatedly to amaze me to this day. Who would have thought that profit-maximizing businessmen benefit by hiring discriminated minorities and paying them more than their competitors? Or that the rational pursuit of a partner would lead to a higher proportion of homosexuals living in cities than in the countryside? Economics always offers simple but not simplistic ways of understanding seemingly complex social phenomena, and although economists tend to be well paid, the pursuit of such understanding is really its own reward.

Indeed, economic analysis is a journey towards ever greater breadth and depth. Some of the many people who have played a big part on my trip have been my Uncle and my late Granddad, who among other things introduced me to the stock market when I was little; Professors Gary Becker and Steve Levitt at Chicago, who convinced me that I, too, could do economics for a living when I had doubted it; my Mother, who on many occasions provides a first testing ground for ideas; my friend Tuomas Nurminen, who
offers additional testing ground, and willing and ever-helpful comments; and of course my eminent and ever-approachable thesis supervisors Rune Stenbacka and Topi Miettinen. In addition, administrative staff who have given me help of particularly high value are Helen Malmsten and Barbara Cavonius. What a pleasure to have had access to all these terrific individuals!

An important principle in Economics is not to waste resources. Since I have said all I wanted to say, it is now time to get on with the dissertation.


## CONTENTS

<table>
<thead>
<tr>
<th>Chapter</th>
<th>Title</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Social Economics: An Overview</td>
<td>1</td>
</tr>
<tr>
<td>2</td>
<td>Buildings and Welfare</td>
<td>22</td>
</tr>
<tr>
<td>3</td>
<td>Puzzling Evidence on Voter Turnout</td>
<td>45</td>
</tr>
<tr>
<td>4</td>
<td>Meal Signals</td>
<td>64</td>
</tr>
<tr>
<td>5</td>
<td>Marriage and Biology</td>
<td>81</td>
</tr>
</tbody>
</table>

---

SOCIAL ECONOMICS: AN OVERVIEW

The field of Social Economics studies influences on the individual from others’ preferences or behaviour. Variously termed the analysis of social interactions, peer effects, or (in a narrower sense) identity economics, Social Economics takes an expanded view of the markets traditionally analysed by economists to include influences by peers as incentives to behave in certain ways. It is not some kind of reaction against “anti-social” economics. Ordinary language contains set phrases like “keeping up with the Joneses” or “peer pressure” to reflect what economists might think of as status consumption or image concerns. Fads are the most obvious instance of people quickly imitating others, but the present dissertation sheds light on a number of other areas in which, while not a dominant element, social economics does help explain certain important phenomena.

This introductory chapter to the present dissertation serves as a background and a summary of some of the important developments in the field of social interactions, mainly with a focus on a combination of factors, such as relevance for the remaining chapters of this dissertation, general degree of utilization within the field, fame, and personal preference. Thus, this is no exhaustive background of the entire field, but the present overview serves the purpose of locating the articles of this dissertation in the field.

As will on occasion be evident in the chapters which constitute this dissertation, many sociologists and anthropologists have concerned themselves with issues highly related to social interactions. Although these treatments will figure only in fairly secondary roles in the present dissertation, it may be useful to take a moment at this early stage to highlight various approaches in social science to social interactions.

In empirical work, a great emphasis on statistical methods is common across disciplines, but another common method which is surely alien to economists is what is known as “thick description” (Geertz, 1973, ch. 1). This is a method whereby combinations of observations, interviews, and studies of “dead” sources such as documents or artefacts are to serve as context to whatever story the researcher is trying to tell. When done well, this method eliminates the problem of identification, since so much background is known that the right explanatory mechanism is made obvious.

Notice that laboratory experiments have difficulty in reaching such a standard for analysing social interactions, because mirroring the whole structure of individuals’ social connections in the lab is nigh on impossible. Economists are, however, typically reluctant to use “thick description” (despite advocacy by some prominent members of the
profession like Akerlof – see below), probably in some part because of unfamiliarity, but mainly because of its requirements on their time and questionable generalizability from individual studies. One economic treatment which nevertheless comes fairly close to “thick description” is by Steve Levitt and Sudhir Venkatesh (2001), which follows the lives of 118 individuals who grew up in a Chicago housing project in the 1980’s. The authors find suggestive evidence that youthful involvement in criminal gangs greatly reduces an individual’s long-term income trajectory, compared to non-gang individuals of similar social standing in the housing project.

In terms of theoretical framework, non-economist social scientists depart from concepts such as maximization and equilibrium. However, this does not mean that their descriptions are necessarily incompatible with these economic concepts, and they might inspire economic analyses to break new ground. One sociological perspective of particular importance is that of symbolic interactionism, closely associated with Erving Goffman. Symbolic interactionism centres on the presentation of self in interaction with others, pithily summarized by Goffman in his book *Interaction Ritual* (2005 [1967], p. 3) as “not, then, men and their moments. Rather, moments and their men”. Individuals take lines to do face work in a way such that they optimally present themselves to others, and what lines are used is determined by the social situation.

While the preference relation under this view is evidently completely social (with the concomitant problem of how societal expectations could ever have been formed – see Wrong, 1961), there is an unmistakeable rational-choice element in Goffman’s account, for people are experts at finding the optimal action to handle the demands of society – “lightning-fast calculators”, to wit. Exceptions to this rule are socially harmful, since they break social conventions and upset the social order, but also highly unusual (pp. 40, 44). There is thus some essentially numinous sense or “meaning” floating around Goffman’s writings which determines what social score individuals get by certain lines, but given that people are socialized into this alleged reality, they act rationally in pursuit of the best sustainable face they can get2. Similar views of pedestrians’ strategies to navigate dangerous neighbourhoods – including the strategic use of grunting and when to switch to the other side of the street, i.e. tactics which mark one’s social position and relation to potential thugs – may be found in Anderson (1990).

2 Goffman aficionados might object that he was less inclined to construct arguments consistent with rational choice in other contexts. For instance, in ‘The Nature of Deference and Demeanor’ (also in *Interaction Ritual*), he writes of rules of conduct, which hold society together, as “recommended” because they are “suitable or just” (pp. 48, 90). Thus, individuals obey them even if it means self-sacrifice. This objection is a fair point, but note that there are ideas consistent with maximization which imply a trade-off between private and social components of the utility function – see, for instance, Akerlof (1997), dealt with in Section III below.
Robert K. Merton is otherwise likely to be the most famous sociologist to talk about the effects of peers and their preferences on behaviour (he is also the father of Robert C. Merton). In work with Alice Kitt, reference groups are salient for individuals because individuals like to associate with people of similar characteristics or with people whom they aspire to be like (Merton and Kitt, 1950, pp. 87-88). The issue of what other individuals belong to one’s reference group is an important one in general and will be dealt with again below.

This non-exhaustive introduction of some of the background in social science, both within and outwith economics, of the present object of study sufficiently shows that social economics deals broadly with peer influences in a way consistent with general economic theory. The remainder of this introductory chapter deals more closely with economic treatments of social interactions, concluding with a summary placing the essays of the present dissertation into perspective.
I - APPROACHES BY PRICE

Although the domain of social economics has for a long time been considered the purview of other social scientists, early economists did pay a great deal of attention to social interactions. For example, Jeremy Bentham (1789) counted the “pleasure of a good name” as one of maybe fifteen different basic human pleasures. However, the formalization of consumer theory brought with it a demand for tractability of utility functions and – as may happen with many novelties – researchers explored the depths (the intensive margin) before considering the width (the extensive margin), with the field of social interactions consequently leading a more slumbering existence in the meantime, albeit with several noteworthy exceptions.

Thus, Thorstein Veblen, a member of the Economics Department of the University of Chicago, is credited with coining the term ‘conspicuous consumption’ in his Theory of the Leisure Class (1899 [1994]). Later approaches by people like James Duesenberry (1949) and Robert Pollak (1976) testify that whatever dominance sociologists have had in the field of social interactions, they have not had a monopoly. These latter two works are examples of introductions of formal models with interdependent preferences, essentially models where an individual’s utility depends not just on his own consumption but also on how it compared to that of his peers.

The most significant contribution to the study of social interactions by economists, however, must surely be the various areas examined by Gary Becker. Social interactions are obviously integral to his analysis of discrimination (1971 [1957]), in which the characteristics (racial, sexual, or whatever) of other persons, beyond their productivity, influence how much one is willing to pay to deal with those persons. Social interactions likewise take centre stage in decisions regarding whom to marry, or fertility, such as when parents choose to have fewer children and perhaps develop deeper connections with them than they might, had they chosen to have more numerous children (Becker, 1973; 1991). Elsewhere, Kandel and Lazear (1992) show how the free-rider problem can be overcome in business partnerships of size with the help of peer pressure.

More generally, aspects of other persons figure in the agent’s utility function when he cares about others beside himself (as manifested, for instance, in “love” or “hate”). Such concerns are behind the reasoning of why members of a family might maximize the general familial welfare rather than “free ride” and care only about themselves, or why people give charitable contributions (Becker, 1965, 1974). With Kevin Murphy and others, Becker (2000) collected treatments of peer effects ranging from why profit-maximizing crowded popular restaurants do not raise their prices, to why seemingly
esoteric books might on occasion climb best-seller lists, or people's valuations of antiques, all of which may be treated as instances in which individuals derive utility in part from partaking in the consumption of something they know is popular.

Social influences are treated as complements in a utility function of the following general form:

\[ U = U(x, y; S), \]

where \( x \) and \( y \) are goods which the individual may consume, with the important addition that one of them, say \( x \), yields different amounts of utility depending on the level of the non-choice variable \( S \), which captures some average of what the individual's peers consume of \( x \). Whereas in the more “traditional” economic approach that ignores social interactions, there would be no \( S \) and no complementarity between \( S \) and \( x \), social economics assumes that there are goods for which peer consumption changes the amount of utility they yield.

Letting \( y \) be the numeraire and \( p_x \) and \( I \) stand for the price of \( x \) and income, respectively, the individual's budget constraint,

\[ p_x x + y = I, \]

may be used in the well-known derivation of the individual's demand function for \( x \) to obtain (under some fairly harmless assumptions) the following expression:

\[ \frac{dx}{dS} > 0. \]

That is, an exogenous rise in \( S \) raises the marginal utility of \( x \) relative to that of \( y \) (adjusted for price), assuming that greater peer consumption positively impacts own utility (sometimes called a “bandwagon effect”). The inequality would go the other way in the opposite case in which greater peer consumption reduces marginal utility (occasionally referred to as the “snob effect”). In the aggregate, this latter effect would have to die out rather quickly (it is not unlike individuals’ incentives to purchase public goods), but it may more commonly happen when \( S \) is an average of the consumption of some out-group, with whom a group of peers may wish to distance themselves, of which perhaps the best-known example is “acting white” (e.g. not speaking Ebonics and getting good grades) in school (Austen-Smith and Fryer, 2005).

When the consumption of a good or service yields greater utility the more other people consume of it, there is a straightforward “social multiplier” in the consumption of said good or service. So if the crowded restaurant in the foregoing example were to raise the prices of its menu items, even just slightly, the patrons it loses purely because of the
increase in prices would cost the restaurateur still more customers via the social multiplier. These models are a very logical step in considering social influences, since norms are commonly thought of as an average (be it the mean, median, or mode) of peer behaviour, and therefore easily translatable to quantitative terms and made amenable for straightforward empirical analysis.

A widespread phenomenon in models of social interactions (not just in price-based models) is that of inefficient escalation. Commonly a result of the fact that status is often gained by relative position, rather than absolute performance, this phenomenon occurs in nature as well as in human societies. Examples from the latter include things like expensive engagement rings or cars. At low production costs, producers will supply too many people with the product in question, with the result that the more exclusive set of people see no point in purchasing it. Consequently, products will be produced at inefficiently high costs to the extent that the gain in utility depends on the difference in quality (production cost) between one’s own purchase and the purchases of others, rather than on absolute quality.

From the field of biology, probably the most commonly-cited example is that of the peacock’s tail. All peacocks would gain in mobility and comfort if their plumage were significantly smaller, but each individual peacock wishes to distinguish himself to peahens by having a more impressive plumage than those of his peers, leading to inefficient escalation. The trees of the forest would similarly gain as much sunlight if they were all shorter and they would avoid spending resources on unnecessary growth, but each individual tree always gains sunlight by additional growth.

Analytically, these situations are similar to that of the signalling literature in which everyone prefers to be considered a “good” type but only a fraction of everyone truly is “good”. When “bad” types can mimic “good” types, the latter must engage in ever-costlier signalling or fail to distinguish themselves.

In human societies, people can to various extents choose which “pond” to be in. Squash players may or may not be absolutely good at squash, but they are apt to be far superior at it relative to their many non-playing friends. The many endeavours and interests a person can have in this way enables him to choose an apposite activity that enables him to be at the top of his field. This vastly exaggerated account of social status contains an important insight that status is not necessarily a zero-sum game and may be turned into a positive-sum one when a felicitous choice of peers is possible, an issue which will recur below in Section III.
II - NON-PRICE APPROACHES

Traditionally, economists have devoted most of their deepest attention to interactions affecting prices and output; if some persons come to want a product more intensely, the greater demand forces its price to increase, possibly leading other consumers to substitute some other product for the present one. Concepts such as methodological individualism, markets, and marginalism form the bedrock of an indispensable method of analysis, focusing on equilibrium and aided by rational choice (individuals seek to maximize something). A constant in the aforementioned treatments by Becker and associated authors is the focus on price even while analysing peer effects.

Other important analyses of social interactions have been less insistent on the importance of price. Schelling’s (1978) tipping points (or critical mass), for instance, essentially posits that persons can take essentially only three attitudes towards an action: always take it; always abstain; or take it if a sufficient number of others do the same. With a broad interpretation of “action”, this framework can be applied to the predominance on one race in many neighbourhoods (which is likely the best-known application of Schelling’s critical-mass framework); for instance, if one wants to live in a neighbourhood only if at least fraction \( x \) is of the same race and this is a common enough preference (even while \( x \) might take different values), racial segregation will arise spontaneously. In his analysis of collective behaviour such as riots and demonstrations, Granovetter (1978) uses a similar framework.

It is a common attribute of public protests that they involve a great number of protestors. Should some of these individuals wish to commit violent acts, the irate attitudes of the people around them are sure to provide them with protection because the police will not be able to detect who threw a stone or committed some other violent act out of a large number of people. The famously peaceful independence movement of India similarly worked because its protesters were carefully screened to exclude those who wish to act violently as well as those who do so if enough others do it. In such demonstrations, stone throwers will be reported to the police and therefore cannot count on the protection of their peers (Tullock and McKenzie, 1978).

Another approach loosely belonging to this category is that based on networks. For instance, immediate neighbours may carry greater weight in one’s utility function than do dwellers just a few houses over, or pupils may learn better if suitably rearranged in a classroom to minimize opportunities for disruptive chatting or unruly behaviour. The analysis of networks has the power to yield insights of this variety, and has been fruitfully employed, for example, by Glaeser et al. (1996) in asking why crime rates differ so much
between cities. Their treatment has people arrayed in a circle, where a person may be, like in Schelling’s (1978) tipping points, of three kinds: the kind who always abstains from criminal activity, the kind that always engages in it, and the kind that does it if his neighbour does it. When these three types are not arrayed identically across cities, crime rates will differ even their relative sizes do not.

In recent years, many economic approaches to social interactions have begun to pay close attention to the social roles of individuals, an approach known as identity economics (Akerlof and Kranton, 2000; 2010). This approach says essentially that people derive utility in some measure from fitting into their social environment. For example, a person of a certain occupation may be supposed to have certain characteristics and act in specific ways. To be a barrister is to dress in a suit and speak formally; a construction worker is sturdy, male, and may have a colourful job history, very many American police officers are black, etc. When individuals take cues from people who are like them to form a perception of their identity, it may be very upsetting when the building blocks of one’s identity are brought into disarray, say by a barrister who swears casually or by a woman construction worker.

Identities may play an important role also in less banal situations. An astounding empirical regularity in family economics is that husbands earn a majority of their households’ market incomes with a sharp discontinuity just after the fifty-per-cent mark (Bertrand et al., 2015). Associated with this regularity is that those households which violate this rule tend to report more problems than others, and have a higher incidence of divorce. Also consistent with this regularity is that the highest-income females are more frequently unmarried than are the highest-income males, a phenomenon which is reversed at the opposite end of the income distribution (see also Saint-Paul, 2015).

While the literature on intra-household division of labour is capable of predicting a vast array of outcomes such as specialization in market and non-market activities and amount of resources spent on children’s human capital, no theoretical justification for the predominance for males earning at least fifty per cent of market incomes with a sharp discontinuity just after this point can be found. Bertrand and her co-authors (2015) conclude from this empirical investigation that there must be something about male identity that says that the man of the house should bring home most of the household’s market income. Again, deviations from this identity cause disutility.

---

3 One of the chapters in this dissertation, ‘Marriage and Biology’, presents an exception to this statement, however.
III - WHICH REFERENCE GROUP?

Many applications will have a fairly reasonable intuitive sense of who is in a reference group, often based on factors such as geographic proximity, race, or age. However, it is equally clear that reference groups are rarely if ever directly observed and the issue of individuals’ “choosing ponds” can pose serious difficulties for researchers. Urging the economics profession to pay closer attention to interview and ethnographic methods practised by anthropologists and sociologists, George Akerlof (1997) develops a gravity-inspired model of social distance, where the gains from trade are a quotient with social distance (the product of current and expected distance), measured as the difference between an individual’s own preference for a certain good or service and that of his peers, in the denominator, and a gain from trade that is constant for all individuals in the numerator (the usual look of gravity-inspired models). “Trade” is here to be understood broadly, including any kind of social intercourse.

In this framework, if a person’s preference for the good or service departs enough from that of his peers, he chooses another “pond” in which to trade, akin to preferences for public goods and moving costs in the seminal Tiebout model (Tiebout, 1956). However, if one’s intrinsic taste for the good or service is not too different from the norm in one’s network, conformity is preferred so as to increase the gains from exchange. This kind of model can capture pressures faced by starkly different groups such as inner-city gang members who might wish to look for different lives or school kids uncertain of their best friends. The relevance of Akerlof’s urging the utilization of more interpretive methods – such as the aforementioned “thick description” – in addition to the common quantitative ones is that they may allow empirical researchers to examine the decision processes of members of the aforementioned groups with a more complete set of background data.

In a series of laboratory experiments, Lazear et al. (2012) allow subjects to choose in what sort of environment they are to interact with other subjects. They offer their subjects the choice of a fixed amount of money or the opportunity to partake in a setting in which they may come to share some of the money, with variations in how conducive the setting is to sharing. They find that, when subjects could opt out at no cost, the number of sharers is reduced, while subsidized and costly sharing situations increase and reduce entry, respectively, mainly by the least and the most generous subjects. In an important way, this replicates real-life possibilities to choose reference groups. Non-sharers are surely better off without the pressure to share in this experiment, while those who do share may continue to feel good about doing so.
The results obtained by these authors align well with Gary Becker’s observation that most people are inclined to avoid beggars when they see them, but nevertheless give money if they fail to avoid an encounter, because their social environment now expects this behaviour (Becker, 1996, pp. 231-233). Lazear et al. (2012) also note other applications of this kind of thinking, such as the way tourists who visit poor countries may go because they like the low prices or because they like to help the indigenous indigents, or, in a more extreme case, how disaster areas attract both the most and the least humanitarian individuals, the former to help and the latter to plunder. A further application concerns advertising, when thought of as a way of associating a product with a particular image (as in Becker and Murphy, 1993). Many products are offered with distinct images and since not everyone consumes every product different “pools” emerge in which individuals have high status (as they themselves perceive it) compared to all those who buy a different brand or do not buy the product at all.
IV - ECONOMETRIC ISSUES

Results like the aforementioned ones indicate the importance of identifying social effects. Charles Manski (2000, p. 127; also 1993, pp. 532–533) describes three separate reasons for researchers to believe that they observe peer effects. One is contextual (or exogenous) interaction, wherein behaviour varies with the (exogenous) characteristics of one’s peers, such as when one bends slightly when among individuals shorter than oneself, or when one observes religious edicts at a holy site.

A different kind of peer effect is known as endogenous interactions, and describes situations in which individuals behave according to actions (as opposed to characteristics) of the group. Examples of endogenous interactions include voting because others vote, applauding when others do, and drinking or taking drugs at parties where drinking and drug-taking happen. Note that neither exogenous nor endogenous interactions imply that individuals must necessarily seek to conform with the group, because they could just as easily take social cues and wish to stand out from what appears to be normal.

The third category of social interaction is known as correlated effects and is observed when individuals have past experiences in common (such as a bad upbringing) so that, in a manner of speaking, their past is what influences behaviour: there is correlation of individual characteristics. An example would be when a band of rascals get up to something rotten, when individual members do it, not because the others do it, nor because the others are rotten, but because they themselves are bad apples. Depending on what sort of history is considered, this effect could also imply something other than conformity although most reference groups would tend to exhibit more conformity than polarization along most dimensions. Notice that this third category is the only one which leads to spurious observations of peer effects, because the observed individuals have preferences and characteristics such that they would want to behave the way they do with or without their peers.

Given the closeness of the first two terms, exogenous and endogenous interactions are occasionally grouped together under the heading of contagion, while correlated effects are also known as homophily. This is quite common practice among sociologists (see also Shalizi and Thomas, 2011). The problems of disentangling homophily and contagion remain severe, however. Because when information with respect to the histories of the persons observed is lacking, it means that one cannot control for effects which are correlated across individuals.
This problem of identification means that empirical researchers face an uphill struggle of innovation and creativity if their studies are to distinguish successfully between influences and shared preferences among individuals. One way to deal with this issue is to find evidence from laboratory work, as in the aforementioned study on sorting into different sharing situations by Lazear et al. (2012).

Outwith the laboratory, subtle social interactions can interfere with economic experimentation. Thus, Carrell et al. (2013) attempt to devise peer groups to optimize academic achievement of incoming cadets at a US Air Force Academy, with a particular focus on raising scores for low achievers (by maximizing the minimum number of incoming cadets in each squadron). Using data on prior cohorts, they determine the optimal configurations for this purpose. However, they find that the scores of the low achievers actually decline slightly (approximately 0.06 grade points) after their intervention (other cadets retain about the same scores). The authors conclude that the students interact mostly within their own levels of ability, confounding the designs of outside experimenters.

Carrell et al. (2013) is testament to the importance of correct identification. The authors believed that the cadets would behave like the cadets with whom they were made to share living spaces (contagion), but the cadets confounded the authors by seeking out a different pond, evidently based upon their individual characteristics (homophily).
V - THE STATE OF THE INQUIRY

Econometric issues aside, many times a study will, even outside the laboratory, obtain results that lend themselves to a social-economic explanation, distinguishing fairly successfully between contagion and homophily. The study by Carrell et al. would qualify as an example. Similarly rich background data enable the famous “obesity-is-contagious” study by Christakis and Fowler (2007), who exploit a double-reporting of friendship data (subjects reported their friends and close friends, so one could check if those named also reciprocated).

Less demanding and more commonplace examples of social interactions include the absence of races close to 42.195 kilometres long – i.e. the official length of a marathon. Runners who complete that race but run no farther must answer in the negative when asked if they have ever run a marathon. Similarly, the 20,000 kilometre distance is rarely contested while the half marathon is very popular. Correlated effects provide an unsatisfactory answer to the concentration of runners at various distances since it is highly unlikely that groups of runners should concentrate out of personal preferences on the common distances and never in-between. The existence of both the imperial and the metric systems of measurements gives additional force to explanations in which the relative achievements of others figure in individuals’ utility functions.

Likewise, mountaineers favour attempting climbs just over 8,000 metres (Europeans, not Americans, have been the main pioneers of this activity) rather than just under. Data from the Indian Himalayas indicate that Annapurna I (8,095 metres) has been attempted by 145 expeditions while Gyachung Kang (7,952 metres) has only been attempted twelve times, and none of the eight summits of less than 8,000 metres in altitude have been attempted more than 50 times, whereas all but one of the eight summits above 8,000 metres have had at least a hundred attempts (Schneider, 2014).

A different set of examples comes from voting behaviour. Voting behaviour has increasingly frequently been thought to have an important social component. It is rather a long-standing puzzle why people would bother to vote at all, given that elections and referenda typically draw sufficiently many people to make one vote exceedingly unlikely to matter. Instrumental theories of voting posit that people vote in order to achieve tangible results, but the low probability of a vote being pivotal has spawned alternatives such as expressive voting (according to which the voicing of sentiment on a ballot motivates the electorate).

Commonplace utterances such as “not voting means one has no right to complain afterwards” (although palpably untrue) suggest that there exists a certain widespread
impression that one ought to vote, at least if one ever has the slightest inclination to
discuss politics or perhaps societal affairs in general. If one abstains from voting, one
must counter the risk of being looked down upon by one's peers, at least when a sufficient
number of them probably did cast their ballots.

While these pressures have long been suspected, it is mainly in recent years that clever
empirical investigations have been able to document them. DellaVigna et al. (2014) find
that, in a survey of a Chicago neighbourhood, people were willing to forego slight
monetary rewards in order to avoid lying, and the questions they were asked to respond
to concerned their voting behaviour. Similarly, Funk (2010) finds that allowing increased
use of mail-in ballots in Switzerland (and a concomitant reduced pressure on polling
stations on days of referenda) reduced voter turnout because it made it easier to abstain
when one's absence from the polling station of a particular day was less conspicuous.
(The present dissertation contains one essay, 'Puzzling Evidence on Voter Turnout', that
also argues in favour of social incentives to vote.)

Pertaining to the issue of reference groups, Roland Fryer and Steve Levitt (2004)
document the rise in distinctively black-sounding names (such as DeShawn, Infinity, or
Precious) across America since the 1960's. Naming conventions are an important part
of cultural and socioeconomic identities and are therefore essential in individuals'
perceptions of their reference groups. Fryer and Levitt find that naming conventions
between whites and blacks began to differ around the heyday of the Black Power
movement in the 1960's, which emphasized pride in one's ethnic origins. Apart from
names, afro hairstyles and afro-centric clothing such as kaftans or various head wraps,
with colour combinations of the red, yellow, and green (which figure prominently on
most flags of African nations but are otherwise a rare combination), also became more
popular beginning around this time.

Fryer and Levitt find that black names occur most frequently in families that live in
neighbourhoods with few whites, and least frequently where many whites live. The
significance for social economics of what Fryer and Levitt document is brought out when
comparing the different approaches to social interactions which this essay has
mentioned.

Thus, if price played the crucial part in naming conventions, presumably because a
black (white) name helps its bearer in interactions with blacks (whites), then black
names should become more common with increased segregation or as the returns to
ability increase in the labour market (assuming that employers tend to be white).
Although more homogenously black communities (where fewer interactions with whites
may be expected) are more likely to see widespread adoption of black-sounding names,
the returns to ability in the labour market have increased dramatically since the 1960’s, and the abolition of the Jim Crow laws in the South have reduced barriers to integration.

Models other than this price-based one are also tested by the authors, who find that an identity model fails to be rejected by their own data. As the 1960’s witnessed a burgeoning Black Power movement, impressions of what it means to be black changed and one would expect those blacks who live in predominantly black neighbourhoods to be most responsive to this sort of change. Interestingly, in the same article Fryer and Levitt find that names, once other socioeconomic background variables have been controlled for, do not appear to reduce labour-market outcomes. This finding is consistent with the failure of the price-based model to account for the change in naming conventions. One would think that a model based upon prices should be more likely to hold if labour-market outcomes were significantly affected by a person’s name.
VI - THE ESSAYS

This section provides a brief overview of the essays which constitute the remainder of this dissertation. They each discuss an area of potential interest in which social interactions shed light on the phenomenon in question. These are, in turn, applications for welfare (i.e. “dole money”), voting, marriage and marital stability, and dietary habits.

The First Essay: ‘Buildings and Welfare’ is an analysis of a collection of visibility-enhancing characteristics of buildings that house welfare offices. Such characteristics include elevated entrances or distance between building and street, so that entrants must spend more time walking towards the building, thereby leaving themselves visible to their judgemental peers. Because of widespread norms that one should live on one’s own work and not be supported by others, a visit to a welfare office is logically paid more frequently when its building characteristics are more amenable to anonymity, and that is precisely what this article finds. Notice that correlated individual effects do not explain the findings because if they did there would be no sensitivity to the physical attributes of the buildings.

The results of this article come mainly from correlational studies with the relevant building characteristics and several control variables as regressors. To pre-empt potential criticisms that such an approach fails to handle possible correlations between building characteristics and unmeasured socioeconomic variables (in spite of all the controls), this article also looks at one area in which eligibility for welfare payments could be ascertained via an online service, as well as the possibility of applying for related welfare payments – which predict eligibility for welfare in general. The building characteristics have a much stronger effect before this reform than they do after.

Aside from bearing witness to the importance of social interactions, this article informs thinking on apposite policy towards welfare seekers. To the extent that one wishes to reduce the stigmatization of welfare applicants, buildings should be equipped with highly anonymizing entrances. However, if one cares more about the efficient use of caseworkers’ time and wants to minimize the number of haphazard welfare applications, one should prefer buildings which put their entrants on the spot.

The Second Essay: ‘Puzzling Evidence on Voter Turnout’ (forthcoming in Rationality and Society) finds that, contrary to established theories of voter turnout, multiple concurrent referenda attract fewer voters than do stand-alone referenda. The finding is robust across the five countries that are studied and the article proposes that voters turn out less because the different issues that are in play imply that different groups may have
a stake in an outcome. If abstention is punished (socially and informally) within groups that have a stake in an outcome, then non-voters get off the hook more easily when other issues are also being voted on, because social punishments are costly and within groups one will wish to let members of the other group (with a stake in the outcome of another referendum) carry out the punishment.

If this explanation is correct, it follows that a substantial number of people vote mainly because their social environment tells them that that is the thing to do. For stand-alone referenda, this means that groups with stronger ties will more easily determine the voting outcome, than in the case in which people voted for reasons internal to themselves. Whether this enhances or reduces social welfare will vary depending on whether social interactions support the side with the highest willingness to pay for an outcome.

With many referenda held concurrently, social interactions introduce noise on the issues that the typical voter lacks social incentives to vote on, but votes on nevertheless because he is already in the voting booth. Many people suspect that voters are biased against sound economic thinking (Caplan, 2007), and if this is correct then multiple concurrent referenda, by attracting voters who would otherwise be silent, can worsen outcomes (cf. Bracco and Revelli, 2017).

_The Third Essay: ‘Meal Signals’_ attempts to ascertain whether the relatively heavy-set members of dinner parties signal conformity to the thinner members’ greater health consciousness by things like ordering light meals or skipping chips as a side dish. This is a signal because one must forego having a better-tasting meal in order to send it. In fact, this is exactly what is observed in the sample. Among lone diners the heavy-set eat sturdier meals while among group-diners the relatively heavy-set order lighter ones even though if anything heavier individuals otherwise eat more calories. Another finding of this essay is that individuals’ orders have a tendency to reflect the orders of the rest of their groups, in line with Christakis and Fowler (2007).

If individuals of more ample frames reduce their caloric intake because they signal health consciousness, then it could be that many displays ostensibly directed at self-help are in fact purchased mainly because of their signalling value. These would include gym memberships, cook books with healthy recipes, or theoretically even stays at addiction clinics. To the extent that these purchases are made for their signalling value, they do not indicate at all that the purchasers should have problems with addiction. Rather, they may be perfectly happy with what they do but not altogether satisfied with how their social environment perceives them.
The Fourth Essay: ‘Marriage and Biology’, finally, looks at a narrower set of peers (a spouse or potential spouse) than do the other chapters. Since only women bear children, only the wife’s infidelity with another man can result in a householder’s unwittingly spending resources to raise offspring with no shared genetic material. While both spouses surely dislike the other’s infidelity, this biological fact means that the man has more to lose from being cheated on. Consequently, the man has a stronger incentive than does the woman to take steps to avoid it. The essay argues that a useful way of doing that is by having an income that is greater than that of his wife, since he can then withhold consumption opportunities from her and the fear of this lower standard of living induces the wife to remain faithful. These incentives can account for the widespread phenomenon that husbands tend to earn at least fifty per cent of their households’ total market income, and also produce a number of other testable implications.
REFERENCES


Schneider, James, *The Seven Deadly Sins*, Unpublished book manuscript (2014).


BUILDINGS AND WELFARE

Mats Ekman

This article looks at buildings on GoogleMaps' 'StreetView' feature, finding that the approval rate of applications for social assistance is higher in welfare offices with building characteristics that enhance the visibility of entry. A fitting explanation for this finding is that persons looking for social assistance dislike being thought of as 'welfare cases' and apply more conservatively when others can see them, particularly if their chances of approval are low, which raises the approval rate by reducing the denominator. The effects decline in the rate of poverty, suggesting that the self-reliance norm weakens as poverty increases. (JEL classification: C90, H29, I38, J13, R53, Z13)

"They all need buildings to help them along"
- David Byrne, The Talking Heads

---

4 Special thanks to several administrators at welfare authorities across the US, as well as to Alexandra Stanczyk for practical advice on the workings of welfare offices while this project was at an early stage. Professors Ola Andersson, Fredrik Carlsson, John List, Topi Miettinen, Elias Oikarinen, Matti Sarvimäki, Rune Stenbacka, and Barbara Wolfe have also given very useful comments on this paper. The usual disclaimer applies.
I - INTRODUCTION

A strong norm that an individual should be self-reliant and get by on his own efforts induces applicants for social assistance (‘welfare’) to feel shame for failing to do so. Known as ‘welfare stigma’, this kind of opprobrium has long been thought to explain why the number of persons eligible for welfare substantially eclipses the number of claims to it (Moffitt, 1983; Lindbeck et al, 1999; Andrade, 2002; Currie, 2006). Despite this pervasive suspicion, empirical evidence of welfare stigma is very hard to come by; for instance, the will to avoid stigma is hard to disentangle from a reluctance to learn about potentially complex application procedures.

Combining data on acceptance rates for Temporary Assistance for Needy Families (TANF) with observations of characteristics of buildings housing welfare offices obtained from GoogleMaps’ ‘StreetView’ feature, this study finds a higher acceptance-to-applications ratio in offices into which entry is more readily seen – as judged by the prevalence of visibility-enhancing features such as elevated entrances or more space to traverse between a building and its street. This result is substantially weaker when information about applicants’ chances of approval is made more accessible online through upgrades of the welfare authorities’ websites, which reduces the need for an in-person appearance at the welfare office. The result is also weaker in areas with greater concentrations of poverty.

An interpretation of these results which is consistent with the literature on welfare stigma is that people are more reluctant to apply for dole money when their peers can more easily observe it, and that those whose poverty is most severe (and whose applications therefore most likely to result in approval) are least likely not to apply. That is, those in most severe poverty may not be more sensitive to social stigma than people making more haphazard applications. This holds if stigma is a normal bad: the better-off the individual, the less stigma he is willing to countenance. The present approach overcomes aforementioned difficulties with disentangling stigma from many other forces because welfare regimes are the same within states but building characteristics vary.

Peer influence in its various forms has been a burgeoning topic in recent times. The paper that is most closely related to the present one contributes results from a laboratory experiment. In a working paper on welfare stigma, Friedrichsen et al. (2016) test the general knowledge of their subjects, finding a thirty percentage-point reduction in take-up rates of transfers to poor performers when a quiz is public rather than private. The ever-present concerns with contextual variation and external validity create a pressing need for field studies on welfare stigma.

From more qualitative sources, Goffman (1963, p. 3) cites interview evidence that being seen to enter a building has figured even in the mind of a library patron. More tangentially, in an interview in Studs Terkel’s (1972) book Working, a shoe-shine man is described to turn down better-paying work because it involves more visibility than does his present place of work (p. 110). In another short interview, a mother of five on welfare mentions the stigma she encounters when people know she receives dole money (pp. 259-262). While not directly related to building characteristics, it would not seem implausible that being seen factors into the experience of welfare applications.

In other topics, social recognition and peer effects are burgeoning factors in several important phenomena, spanning voting (e.g. Funk, 2010), littering (Dur and Vollaard, 2014), fairness and charitable giving (Dana et al., 2007; Lazear et al., 2012; Krupka and Weber, 2013; Butera and Horn, 2017), and going back to Thorstein Veblen’s (1899) seminal treatment of conspicuous consumption (see also, e.g., Frank, 1985; Samuelson, 2004). The impact of social incentives is also thought to depend negatively on ‘price’; norms are mutable and if it exogenously became much harder to fulfil social expectations, the norms would weaken (Bernheim, 1994; Akerlof, 1997; Becker and
A norm may be internalized or it may come from outside the individual. In the former case, man is said to be ‘socialized’, whereas the latter case is more in tune with methodological individualism, as norms are then more akin to others’ expectations of an individual, which he can choose to disregard (Wrong, 1961; cf. Elster, 1989). Along with internalized norms (e.g., López-Pérez, 2008, 2010; Kimbrough and Vostoknutov, 2016), reputational or self-respect concerns (e.g., Bénabou and Tirole, 2006; Malmendier et al., 2014) and caring about others’ material payoffs (e.g. Becker, 1974; Fehr and Schmidt, 1999) are common approaches to social interaction. Models in psychological game theory in which players receive disutility (“guilt”) from failing to live up to others’ expectations of them, may also be thought of as internalized norms (e.g. Geanakoplos et al., 1989; Rabin, 1993; Baumeister et al., 1994; Dufwenberg, 2002; Battigalli and Dufwenberg, 2007).

If social norms are not (perfectly) internalized, the utility loss due to breaking them is a function of the likelihood that a ‘welfare case’ be seen to enter a welfare office; socialized individuals would plainly abide by a norm whether watched or no. Anonymity allows unsocialized individuals to disregard self-reliance norms with impunity. While social assistance administrations invariably protect applicants’ anonymity very strictly, protection is necessarily incomplete when in-person applications must be made at welfare offices so that others can see who goes in. Although the results of the present study cannot rule out any socialization of norms, they suggest the presence of unsocialized norms.

In addition, this paper accomplishes two main things. Firstly, it contributes a novel method of indirectly observing norms and social interaction effects based on differences in visibility (or anonymity), provided that visibility differences are not correlated with relevant socio-economic features or that felicitous circumstances enable the socio-economic features to be controlled for by apposite empirical design such as regression discontinuity design or differences in differences. Although visibility is integral to the presentation of self, there appears to be no prior studies which utilize visibility-affecting variations in physical structures to examine differences in behaviour.

Secondly, the results of this paper are consistent with theories of welfare stigma and thus suggest that building characteristics influence welfare-seeking decisions, and additionally that welfare stigma is weakly lower where poverty is higher. Therefore, norms do appear to respond to prices, consistent with the aforementioned research on the mutability of norms. If the loss of status from being seen to enter a welfare office diminishes with increased poverty in this way, a local welfare office could come to act as a way for individuals of connecting with one another at sufficiently high levels of poverty, in effect turning aversion into attraction and thereby ‘normalizing’ a ‘welfare culture’ (e.g. Coleman, 1988, Bertrand et al., 2000; cf. Carvalho et al., 2017).

The potential policy relevance of this paper is especially linked to the literature on ‘welfare ordeals’ (e.g. Besley and Coate, 1995; Blumkin et al, 2015), which is concerned with ways of gaining access to welfare payments that are less costly for the ‘truly needy’ than for the merely indolent who seek money for nothing. Shame is a nonpecuniary cost that is plausibly incurred more willingly by those with fewest resources, whose welfare applications are more likely to be approved than those in less dire straits. Visibility-enhancing building characteristics could therefore economize on social workers’ time. Whether this is socially efficient depends on how much public resources are saved in this fashion and how severe are the costs of stigma of ‘deserving’ applicants.

The rest of this paper proceeds as follows. Section II describes the data and gives some institutional background. It also discusses methodological issues of the study. Section III
provides the main results, which are consistent with the literature on social influences on individuals’ behaviour. It also shows some weaker evidence that norms are stronger in higher-income service areas. Section IV offers a brief discussion relating the findings to the complementarities approach to social interactions and highlights some implications which arise from this framework, whereupon Section V concludes.
II – DATA AND METHODS

The mechanism which the present paper proposes should now be clear; because of a widespread unwillingness to be considered a ‘welfare case’, people apply for welfare more conservatively when others are likelier to see it, driving up the acceptance ratio as those most certain of approval are least likely not to apply, ceteris paribus. The relevant visibility variables are the seven ones described below. Each such characteristic of buildings housing welfare offices is assigned a value of one if it increases the visibility of entry and nought otherwise, as determined by their pictures obtained from the ‘StreetView’ feature of GoogleMaps (N=536). No research assistance was used in collecting the data and the coding was done before checking the applications data and control variables.

The building characteristics are the following (only these characteristics have ever been considered), all in binary form:

1. *Elevated entrance:* An elevated entrance, such as a set of steps or a natural slope leading to the door, makes entry more visible by raising entrants above their peers on the street as they walk in. One is more easily seen and may therefore be more reluctant to go in.

2. *Corner entrance:* An entrance on a corner can reduce the social stigma that entrants face since fewer peers may observe those who enter. It is relatively easy to ‘sneak in’ unsuspectingly, as potential onlookers will believe that one has simply rounded the corner.

3. *Busy street:* A well-trafficked street naturally discourages people to enter any building that lowers one’s status, since more people can see where one goes. Beyond some point, however, a crowded street may have an anonymizing effect, which will be evidenced below.

4. *Space between building and street:* The greater the distance one must traverse between the street and the building, the more one is discouraged from entry, since there is more time during which one may be observed.

5. *Many clear windows:* The more clear windows a building has, the more people can see who is inside it. Entrants pondering this fact may be reluctant to enter many-windowed buildings.

6. *Shared building:* If the building which houses a welfare office is also home to other functions, pedestrians cannot infer that an entrant is a welfare applicant, which should tend to increase applications.

7. *Large sign:* While large signs may theoretically increase the number of welfare applications by letting welfare-seekers know where to apply, online resources, phone books and the like make this effect unlikely (indeed, the fact that offices have been located online for this study is evidence that finding them is not too difficult, even without local presence). Instead, a large sign is apt to inform onlookers that the entrant is out to apply for welfare payments.

Except for *Shared building* and *Corner entrance*, these characteristics are continuous, and coding them in binary form necessarily involves grey areas. Unfortunately, there is no feasible way in which GoogleMaps can be used to determine, for instance, the difference in altitude between a door and a street to measure elevation. However, there is no reason to suspect systematic errors of mistaking more strongly present characteristics (such as an unusually large and plentiful windows on a building) for absent or insignificantly present characteristics, or vice versa.
Combinations of these characteristics may complicate outcomes. This only happens six times in the sample and has no impact on the results, but mention must be made to fully account for the method. For instance, an entrance may be located on a corner, but still be elevated, in which case the elevation ruins the corner effect, since nobody who rounds a corner in a normal fashion would simultaneously appear to be on an ascending trajectory, although if the elevation is reachable from both sides of the corner one might, since it may then be a shortcut round the corner.

A related case is the location of a corner. If a street bends in a corner, sans contact with another street, that is the optimal corner entrance for a welfare applicant who wants to avoid stigma. This propitious corner situation would have street fractions A and B removed, as in Figure I; in which the welfare applicant is approaching the welfare office (arrow), marked X. If A and B came back, the corner advantage would lessen; although views of the welfare applicant from street fractions C and D remain obscured by the corner, the views from street fractions A and B are not. However, since the entrant appears but briefly to persons on fraction A, a corner entrance is sure to give some reduction in visibility.

A final complicating factor is that the presence of many clear windows is not visibility-enhancing when there is much space between the street and the building, simply due to the added distance. The windows variable essentially just means ability to look inside.

Six states contribute data: Alabama (65), Georgia (148), Missouri (101), North Carolina (92), Oklahoma (71), and Oregon (59). In all of these states bar Oregon, welfare authorities pool application data by county. This necessitates the exclusion of some of the bigger cities, which occupy one county but run several welfare offices whose applicants consequently cannot be tied to a particular building. However, cities like Savannah, GA, Birmingham, AL, and of course all the big cities of Oregon remain in the sample.

A total of 48 service areas are not in the sample because of the issue of multiple offices, or because GoogleMaps could not find the offices (usually because the area had not been visited by the Google Car; sometimes because the address could not be found). Data on the number of total applications for Temporary Assistance for Needy Families (TANF), as well as approved applications, are available for the remaining 536 welfare offices. The North Carolinian data come via the Jordan Institute for Families at the School of Social Work of the University of North Carolina, Chapel Hill; otherwise all data are direct from state Departments of Human Services. The application data are aggregated over a calendar year, ranging from 2010 to 2014, depending on availability.

According to the official guiding principles that Departments of Human Services use when selecting a building to house a welfare office (frequently available upon request), offices are chosen to be within easy reach for people in their service areas, to offer adequate and clean office space, and bathroom facilities, as well as attractive rent. There is no indication that offices ever be chosen for their visibility-impacting building characteristics.

The Center on Budget and Policy Priorities (2015) reports that the share of families under the poverty line that receives TANF payments varies from approximately forty per cent in Oregon, to seven in Georgia. In-between are Missouri (29), Alabama (17), Oklahoma (10), and North Carolina (9). The US average for the period studied is about 25. The details of TANF provision also vary between states, but for the purposes of the present paper it suffices to note that the programme is aimed at providing cash for indigent families.
families with dependant children under the age of eighteen (or nineteen if still in high school), and that the shame of being seen to apply for it is concentrated to the actual application (unlike, say, food stamps, where there may be stigma each time they are used). Lifetime limits on the number of months these payments may be made to the applicant (often 60 months), and conditions on (eventually) finding employment for receiving aid, also vary.

If welfare offices handled many things that most people faced no social opprobrium for doing, the proposed mechanism would not work since onlookers would be unable to infer that entrants want social assistance of some kind. With one insignificant exception5,

---

5 The exception is adoptions, but this is unlikely to upset the visibility-opprobrium mechanism since adoptions are much rarer than is reception of TANF benefits. In a typical year, around 120,000 children are

---

Table I: Summary Statistics

<table>
<thead>
<tr>
<th></th>
<th>N=536</th>
<th>AL (65)</th>
<th>GA (148)</th>
<th>MO (101)</th>
<th>NC (92)</th>
<th>OK (71)</th>
<th>OR (59)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Busy street</td>
<td>266 (50%)</td>
<td>25 (38%)</td>
<td>67 (45%)</td>
<td>51 (50%)</td>
<td>37 (40%)</td>
<td>47 (66%)</td>
<td>39 (66%)</td>
</tr>
<tr>
<td>Windows</td>
<td>62 (12%)</td>
<td>5 (8%)</td>
<td>14 (9%)</td>
<td>12 (12%)</td>
<td>10 (11%)</td>
<td>11 (15%)</td>
<td>10 (17%)</td>
</tr>
<tr>
<td>Space btw street and bldg.</td>
<td>367 (68%)</td>
<td>44 (68%)</td>
<td>82 (55%)</td>
<td>82 (81%)</td>
<td>77 (84%)</td>
<td>48 (68%)</td>
<td>34 (58%)</td>
</tr>
<tr>
<td>No corner entrance</td>
<td>498 (93%)</td>
<td>62 (95%)</td>
<td>124 (84%)</td>
<td>97 (96%)</td>
<td>90 (98%)</td>
<td>68 (96%)</td>
<td>57 (97%)</td>
</tr>
<tr>
<td>Elevation</td>
<td>114 (21%)</td>
<td>16 (25%)</td>
<td>37 (25%)</td>
<td>29 (29%)</td>
<td>16 (17%)</td>
<td>10 (14%)</td>
<td>6 (10%)</td>
</tr>
<tr>
<td>No shared building</td>
<td>457 (85%)</td>
<td>62 (95%)</td>
<td>123 (83%)</td>
<td>88 (87%)</td>
<td>73 (79%)</td>
<td>63 (89%)</td>
<td>48 (81%)</td>
</tr>
<tr>
<td>Large sign</td>
<td>132 (25%)</td>
<td>17 (26%)</td>
<td>62 (42%)</td>
<td>9 (9%)</td>
<td>27 (29%)</td>
<td>11 (15%)</td>
<td>6 (10%)</td>
</tr>
<tr>
<td>County mean rate of poverty</td>
<td>0.121</td>
<td>0.127</td>
<td>0.166</td>
<td>0.093</td>
<td>0.109</td>
<td>0.099</td>
<td>0.091</td>
</tr>
<tr>
<td>County median rate of poverty</td>
<td>0.109</td>
<td>0.120</td>
<td>0.168</td>
<td>0.091</td>
<td>0.104</td>
<td>0.091</td>
<td>0.094</td>
</tr>
<tr>
<td>County median household income</td>
<td>$39,872</td>
<td>$37,277</td>
<td>$37,349</td>
<td>$39,519</td>
<td>$40,180</td>
<td>$41,881</td>
<td>$46,885</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Total applications</td>
<td>146,457</td>
<td>28,955</td>
<td>46,354</td>
<td>3,007</td>
<td>47,684</td>
<td>11,785</td>
<td>8,672</td>
</tr>
<tr>
<td>Mean acceptance-to applications ratio</td>
<td>0.471</td>
<td>0.298</td>
<td>0.203</td>
<td>0.641</td>
<td>0.677</td>
<td>0.324</td>
<td>0.894</td>
</tr>
<tr>
<td>Median acceptance-to-applications ratio</td>
<td>0.449</td>
<td>0.281</td>
<td>0.194</td>
<td>0.647</td>
<td>0.669</td>
<td>0.323</td>
<td>0.904</td>
</tr>
<tr>
<td>Standard deviation</td>
<td>0.269</td>
<td>0.141</td>
<td>0.081</td>
<td>0.090</td>
<td>0.144</td>
<td>0.152</td>
<td>0.090</td>
</tr>
<tr>
<td>Range</td>
<td>0 – 1</td>
<td>0.06 – 0.91</td>
<td>0.03 – 0.52</td>
<td>0.33 – 1</td>
<td>0.35 – 1</td>
<td>0.06 – 0.69</td>
<td>0.40 – 1</td>
</tr>
</tbody>
</table>

Correlations

<table>
<thead>
<tr>
<th></th>
<th>Busy street</th>
<th>Win.</th>
<th>Space</th>
<th>No corner</th>
<th>Elevat.</th>
<th>No sh’d bldg</th>
<th>Large sign</th>
<th>Poverty rate</th>
<th>Dist.</th>
<th>Share black</th>
</tr>
</thead>
<tbody>
<tr>
<td>Busy street</td>
<td>1.000</td>
<td>0.178</td>
<td>0.003</td>
<td>0.027</td>
<td>0.027</td>
<td>-0.014</td>
<td>-0.052</td>
<td>-0.081</td>
<td>0.006</td>
<td>-0.094</td>
</tr>
<tr>
<td>Windows</td>
<td>1.000</td>
<td>-0.171</td>
<td>-0.036</td>
<td>-0.004</td>
<td>-0.012</td>
<td>-0.012</td>
<td>-0.039</td>
<td>-0.031</td>
<td>-0.027</td>
<td>0.005</td>
</tr>
<tr>
<td>Space btw st. and bldg.</td>
<td>1.000</td>
<td>0.189</td>
<td>0.059</td>
<td>0.056</td>
<td>-0.056</td>
<td>-0.056</td>
<td>0.066</td>
<td>-0.062</td>
<td>0.066</td>
<td>-0.030</td>
</tr>
<tr>
<td>No corner entr.</td>
<td>1.000</td>
<td>0.038</td>
<td>-0.014</td>
<td>-0.012</td>
<td>-0.084</td>
<td>0.046</td>
<td>-0.056</td>
<td>-0.084</td>
<td>0.046</td>
<td>-0.056</td>
</tr>
<tr>
<td>Elevation</td>
<td>1.000</td>
<td>0.002</td>
<td>0.041</td>
<td>0.031</td>
<td>-0.110</td>
<td>-0.001</td>
<td>0.043</td>
<td>-0.001</td>
<td>0.181</td>
<td>-0.000</td>
</tr>
<tr>
<td>No sh’d bldg</td>
<td>1.000</td>
<td>0.153</td>
<td>0.078</td>
<td>-0.001</td>
<td>0.043</td>
<td>-0.048</td>
<td>0.181</td>
<td>-0.048</td>
<td>0.181</td>
<td>-0.000</td>
</tr>
<tr>
<td>Large sign</td>
<td>1.000</td>
<td>0.015</td>
<td>0.097</td>
<td>0.072</td>
<td>0.043</td>
<td>0.042</td>
<td>0.043</td>
<td>0.042</td>
<td>0.043</td>
<td>-0.000</td>
</tr>
<tr>
<td>Poverty rate</td>
<td>1.000</td>
<td>-0.120</td>
<td>0.652</td>
<td>0.043</td>
<td>0.042</td>
<td>0.043</td>
<td>0.043</td>
<td>0.043</td>
<td>0.043</td>
<td>-0.000</td>
</tr>
<tr>
<td>Distance</td>
<td>1.000</td>
<td>-0.042</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Share black</td>
<td>1.000</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Note: The (somewhat abbreviated) building characteristics are the ones defined in Section II of this article and are all binary variables. In the lower part of this table that shows correlations, Distance is number of kilometres between the building and the point GoogleMaps regards as the town centre. Share black is the fraction of the population in the county of the welfare office that is black.
the sampled welfare offices provide only such services as might be desired by those who do not rely on work for sustenance. The specifics vary by state, but often include things like food stamps, Medicaid or energy assistance (electric bills). In some states, TANF applicants can simultaneously also apply for an additional welfare payment and being approved for one counts as an approval in those states (but TANF must have been among the applications submitted). These differences are behind some of the big variations in average approval rates between states seen in Table I. Official requirements to receive TANF do not vary between counties.

The sampled states differ from other states in that they do not allow TANF applications to be made online (except for the upgrades to the websites of the welfare authorities in North Carolina and in Oklahoma, mentioned in the introduction), or in other easy ways that side-step the applicant having to enter the welfare office in person. In recent times, several states have introduced such possibilities, which may mean that visibility-enhancing building characteristics become irrelevant for applicants concerned with the presentation of self, if they disvalue the cumbersomeness of reading through online applications more than they disvalue stigma.

The dependant variable is the ratio between the number of accepted TANF applications, and the total number of TANF applications made to a welfare office. The rationale for using this construction is that, if visibility-enhancing building characteristics reduce the propensity to apply for welfare benefits, mainly those individuals most likely to have their applications accepted will risk being seen to apply, causing the ratio to increase through a reduction in its denominator. The main regression model is of the following form:

\[ \text{Approval ratio} = \text{Constant} + \beta \times \text{visibility characteristics} + \gamma \times \text{controls} + \text{error} \]

The choice of dependant variable allows easier comparisons between large- and small-population service areas and limits problems due to the aforementioned relative rarity of TANF families (which makes the poverty rate and population very blunt controls), and to the possibility of caseworker heterogeneity across offices and outliers in high-application service areas.

The control variable distance from centre is constructed by asking GoogleMaps for the directions from the town’s centre to the address of the welfare office. It was mentioned above how welfare offices try to locate to ease access and the median distance from the town centre that this measure obtains is a low 0.9 miles (1.44 kilometres). Considering this method treats the centre as a single point, one may say that the median office is located centrally. Another control, fraction in town, is the share of the population in a service area living in the town of the welfare office (mean 26 per cent). Since welfare offices are commonly located centrally, fraction in town captures proximity to welfare offices for applicants who do not reside in the same town.

adopted in the US (USDHHS, 2004), a figure far lower than the aforementioned TANF caseload, even when doubling it to account for two adoptive parents. It is much more common to enter a welfare office wanting social assistance than a child, and TANF is only one of many welfare programmes. Moreover, it may be that reluctance to enter a welfare office is due to the risk of being seen by an acquaintance, in which case one cannot credibly claim to be considering adopting a child.

6 Even if visibility-enhancing building characteristics turned away some applicants certain of being approved (perhaps because the expected payment is not large enough to warrant social disapproval), the approval rate nevertheless declines. If two welfare offices are identical and approve some fraction \( x/y < 1 \) of their clients for TANF payments, and then for some exogenous reason one of them adds a few steps to its entrance (say) so that one of its clients, who would have been approved, chooses not to apply, its new approval rate is \( (x-1)/(y-1) \), which is less than \( x/y \) since \( x < y \) (multiply by both denominators to get \( xy - y < xy - x \)). Of course, if the reduction comes mainly from haphazard applicants, the ratio declines by more.
Other control variables include the percentage of families with dependant children whom the Census Bureau deems poor, county population, and median household income, all available from the Census Bureau. Neighbourhood income data (i.e. narrower than income measures at the county level – often as close as just the nearest few blocks) are obtained for the state of Georgia from the online resource City-Data (www.city-data.com).
### Table II: OLS Results with TANF Approvals-to-Applications Ratio as Dependant Variable

<table>
<thead>
<tr>
<th></th>
<th>(1)</th>
<th>(2)</th>
<th>(3)</th>
<th>(4)</th>
<th>(5)</th>
<th>(6)</th>
<th>(7)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Busy street</td>
<td>0.029</td>
<td>0.003</td>
<td>0.003</td>
<td>-1.2*10^-4</td>
<td>-0.001</td>
<td>–</td>
<td>–</td>
</tr>
<tr>
<td></td>
<td>(1.252)</td>
<td>(0.251)</td>
<td>(0.285)</td>
<td>(0.012)</td>
<td>(0.143)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Windows</td>
<td>0.091</td>
<td>0.034</td>
<td>0.034</td>
<td>0.037</td>
<td>0.038</td>
<td>–</td>
<td>–</td>
</tr>
<tr>
<td></td>
<td>(2.368)</td>
<td>(1.952)</td>
<td>(1.982)</td>
<td>(2.208)</td>
<td>(2.348)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Elevated entrance</td>
<td>-0.032</td>
<td>0.010</td>
<td>0.010</td>
<td>0.012</td>
<td>0.011</td>
<td>0.012</td>
<td>0.012</td>
</tr>
<tr>
<td></td>
<td>(1.196)</td>
<td>(0.752)</td>
<td>(0.754)</td>
<td>(0.942)</td>
<td>(0.870)</td>
<td>(0.966)</td>
<td>(0.937)</td>
</tr>
<tr>
<td>No corner entrance</td>
<td>0.164</td>
<td>0.027</td>
<td>0.028</td>
<td>0.025</td>
<td>0.029</td>
<td>0.024</td>
<td>0.023</td>
</tr>
<tr>
<td></td>
<td>(3.967)</td>
<td>(1.747)</td>
<td>(1.774)</td>
<td>(1.680)</td>
<td>(1.945)</td>
<td>(1.614)</td>
<td>(1.561)</td>
</tr>
<tr>
<td>Space</td>
<td>0.077</td>
<td>0.020</td>
<td>0.020</td>
<td>0.022</td>
<td>0.024</td>
<td>0.018</td>
<td>0.017</td>
</tr>
<tr>
<td></td>
<td>(3.010)</td>
<td>(1.929)</td>
<td>(1.914)</td>
<td>(2.149)</td>
<td>(2.268)</td>
<td>(1.802)</td>
<td>(1.716)</td>
</tr>
<tr>
<td>No shared building</td>
<td>0.001</td>
<td>0.013</td>
<td>0.013</td>
<td>0.018</td>
<td>0.025</td>
<td>0.018</td>
<td>0.017</td>
</tr>
<tr>
<td></td>
<td>(0.039)</td>
<td>(1.014)</td>
<td>(0.965)</td>
<td>(1.322)</td>
<td>(1.774)</td>
<td>(1.272)</td>
<td>(1.226)</td>
</tr>
<tr>
<td>Large sign</td>
<td>-0.089</td>
<td>0.018</td>
<td>0.019</td>
<td>0.018</td>
<td>0.019</td>
<td>0.020</td>
<td>0.022</td>
</tr>
<tr>
<td></td>
<td>(3.634)</td>
<td>(1.571)</td>
<td>(1.636)</td>
<td>(1.596)</td>
<td>(1.648)</td>
<td>(1.761)</td>
<td>(1.881)</td>
</tr>
<tr>
<td>Fraction in Town</td>
<td>–</td>
<td>-0.007</td>
<td>-0.005</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>–</td>
</tr>
<tr>
<td></td>
<td>(0.384)</td>
<td>(0.293)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Distance fr. Centre</td>
<td>–</td>
<td>-0.002</td>
<td>-0.001</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>–</td>
</tr>
<tr>
<td></td>
<td>(0.712)</td>
<td>(0.428)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Poverty rate</td>
<td>–</td>
<td>0.564</td>
<td>0.545</td>
<td>0.553</td>
<td>0.559</td>
<td>0.531</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(4.107)</td>
<td>(3.849)</td>
<td>(4.161)</td>
<td>(4.190)</td>
<td>(3.891)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>County population</td>
<td>–</td>
<td>–</td>
<td>-3.4^a</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>-5.7*10^a</td>
</tr>
<tr>
<td></td>
<td>(0.705)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>(1.349)</td>
<td></td>
</tr>
<tr>
<td>Alabama Dummy</td>
<td>–</td>
<td>-0.384</td>
<td>-0.384</td>
<td>-0.386</td>
<td>-0.377</td>
<td>-0.388</td>
<td>-0.388</td>
</tr>
<tr>
<td>Georgia Dummy</td>
<td>–</td>
<td>-0.500</td>
<td>-0.501</td>
<td>-0.500</td>
<td>-0.467</td>
<td>-0.502</td>
<td>-0.504</td>
</tr>
<tr>
<td></td>
<td>(27.570)</td>
<td>(27.464)</td>
<td>(27.575)</td>
<td>(27.130)</td>
<td>(27.911)</td>
<td>(27.816)</td>
<td></td>
</tr>
<tr>
<td>Missouri Dummy</td>
<td>–</td>
<td>-0.023</td>
<td>-0.025</td>
<td>-0.025</td>
<td>-0.034</td>
<td>-0.024</td>
<td>-0.027</td>
</tr>
<tr>
<td></td>
<td>(1.257)</td>
<td>(1.324)</td>
<td>(1.369)</td>
<td>(1.891)</td>
<td>(1.327)</td>
<td>(1.456)</td>
<td></td>
</tr>
<tr>
<td>Oklahoma Dummy</td>
<td>–</td>
<td>-0.342</td>
<td>-0.345</td>
<td>-0.343</td>
<td>-0.349</td>
<td>-0.342</td>
<td>-0.346</td>
</tr>
<tr>
<td>Oregon Dummy</td>
<td>–</td>
<td>0.248</td>
<td>0.249</td>
<td>0.236</td>
<td>0.226</td>
<td>0.237</td>
<td>0.245</td>
</tr>
<tr>
<td>Constant</td>
<td>0.268</td>
<td>0.555</td>
<td>0.558</td>
<td>0.548</td>
<td>0.598</td>
<td>0.556</td>
<td>0.566</td>
</tr>
<tr>
<td>F-test, p-value</td>
<td>0.000</td>
<td>0.041</td>
<td>0.038</td>
<td>0.015</td>
<td>0.005</td>
<td>0.039</td>
<td>0.042</td>
</tr>
<tr>
<td>R^2</td>
<td>0.084</td>
<td>0.824</td>
<td>0.824</td>
<td>0.830</td>
<td>0.824</td>
<td>0.829</td>
<td>0.829</td>
</tr>
</tbody>
</table>

Notes: The TANF Ratio is computed by dividing the number of approved TANF applications by the number of applications, so it cannot leave the 0-1 range. All columns report estimates with this ratio as the dependent variable. Thus, the coefficients should be interpreted as the change in the rate of approvals from the presence of one of the binary variables, e.g. a busy street increases the percentage of approved applications by 2.9 points in column (1). The variable Fraction in town is the share of the service area’s population living in the same town as the welfare office. Poverty rate goes from nought to one, and measures the fraction of household with children under age 18 living below the Census Bureau’s poverty line. Distance from centre is a measure of how far away from the city centre the welfare office is located. Except columns (2) and (3) (N=520), all have a sample of 536. The F-test reports the p-value of a test of joint significance of all the building characteristics. Heteroscedasticity-robust t-statistics are given within parentheses. The Bonferroni-adjusted t-tests’ critical values for statistical significance are approximately 2.46 (ten per cent), 2.70 (five per cent), and 3.20 (one per cent), depending very slightly on sample sizes. This adjustment corrects for testing seven building characteristics.
III – RESULTS AND INTERPRETATION

Table II presents some initial findings of this study, with the TANF ratio as regressand, and building characteristics and controls as regressors. Additional regression models with variables capturing the ethnic composition of service areas, population density in the town of the welfare office, and shares of single-headed families (male and female) are omitted to save space but change nothing for the building variables.

Some variables may be particularly prone to coding error, which is why columns (6), (7), and (8) estimate regression models without the variables busy street and windows. Ascertaining the clearness of windows or the busyness of streets is sometimes difficult on GoogleMaps; the other variables are much more easily determined by simple eyeballing.

The fraction in town variable (i.e. the share of the population in a service area living in the town of the welfare office) captures reliance on cars or public transportation (data on which are only sporadically available on GoogleMaps) to reach the welfare office. Its coefficient is small and not statistically significant. Distance from centre, too, is insignificant, indicating that the aim of Human Services departments to locate welfare offices within easy reach of their clientele is successful.

Given that the individual building variables are all binary while the actual characteristics lie on a continuum, none of the dummies is likely to have a constant effect and the signs are more important than the exact magnitudes. All the building variables generally have the hypothesized sign, especially with state controls (recall from Table I that approval rates differ widely by state). In no specification is a building characteristic statistically significant at even the ten per cent level with a Bonferroni correction.

However, it is important to note that the hypothesis is whether visibility on the whole impacts applicant behaviour rather than individual visibility-enhancing characteristics, so that focus should be on the joint statistical significance of the building characteristics. Therefore, an F-test on joint significance is used. In none of the regression equations in Table II is its associated p-value above 0.042, indicating a statistically significant positive correlation between visibility and conservative application behaviour.

Since the TANF ratio ranges from zero to one, the natural social-stigma interpretation is that visibility reduces the number of relatively haphazard applications, thereby shrinking the denominator and raising the acceptance rate. For instance, the absence of a corner entrance is estimated to increase the acceptance rate by about 2.5 per cent. Table I shows that the mean acceptance rate in the sample is 47.1 per cent. At this point, more space (say, a lengthening of the pathway to a welfare office), which raises the approval rate by about two per cent, reduces the number of applications by roughly four per cent at the margin (2/0.491 ≈ 4.07).

The introduction mentions how one building variable, busy street, may work differently in high- and low-population areas. Welfare applicants might feel more anonymous in more populous counties because in sufficiently large crowds one must focus on getting through rather than on recognition of individuals. Possibly this creates a sense of anonymity; everyone might soon know if one has been seen to enter a welfare office in a small town, but in a large town no-one might notice. Dividing the sample at the median county population, 28,484, and regressing the standard model on the high- and low-population subsamples suggests that busy streets do indeed anonymize welfare applicants beyond some point, as Table III reports in columns (3) to (6). Interestingly, this table also suggests that most of the effect from elevated entrances and buildings with large signs comes from the high-population half of the sample.

---

7 However, a number of things about streets are easy to see. E.g., cul-de-sacs are always assigned a value of nought. GoogleMaps also colour streets differently according to traffic, which greatly helps with coding.
Notice that, even though the other building characteristics do not stay the same in the low- and high-population subsamples, only the busy street variable changes sign. The p-values on the F-tests also remain fairly strong, never exceeding 0.024 in Table III. Columns (1) and (2) provide regressions of a model where busy street is interacted with county population, which yield similar results. In these interactions, differentiating with respect to busy street and calculating the population at which a busy street becomes crowded enough to yield anonymity is about 76,000 (0.016/0.00000021).

Table III: Anonymizing Effects of Higher Population on Busy Streets

<table>
<thead>
<tr>
<th></th>
<th>(1)</th>
<th>(2)</th>
<th>(3)</th>
<th>(4)</th>
<th>(5)</th>
<th>(6)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Busy street</td>
<td>0.016</td>
<td>0.016</td>
<td>-0.026</td>
<td>-0.029</td>
<td>-0.026</td>
<td>0.029</td>
</tr>
<tr>
<td></td>
<td>(1.405)</td>
<td>(1.417)</td>
<td>(1.984)</td>
<td>(1.905)</td>
<td>(1.980)</td>
<td>(1.940)</td>
</tr>
<tr>
<td>Busy street × population</td>
<td>-2.1*10⁻⁷</td>
<td>-2.1*10⁻⁷</td>
<td>0.028</td>
<td>0.032</td>
<td>0.031</td>
<td>0.032</td>
</tr>
<tr>
<td></td>
<td>(3.018)</td>
<td>(3.005)</td>
<td>(1.147)</td>
<td>(1.408)</td>
<td>(1.204)</td>
<td>(1.399)</td>
</tr>
<tr>
<td>Windows</td>
<td>0.038</td>
<td>0.039</td>
<td>0.020</td>
<td>0.002</td>
<td>0.020</td>
<td>0.003</td>
</tr>
<tr>
<td></td>
<td>(2.283)</td>
<td>(2.318)</td>
<td>(1.232)</td>
<td>(0.131)</td>
<td>(1.209)</td>
<td>(0.162)</td>
</tr>
<tr>
<td>Elevated</td>
<td>0.008</td>
<td>0.008</td>
<td>0.038</td>
<td>0.009</td>
<td>0.036</td>
<td>0.009</td>
</tr>
<tr>
<td>Entrance</td>
<td>0.678</td>
<td>0.634</td>
<td>1.825</td>
<td>0.450</td>
<td>1.692</td>
<td>0.476</td>
</tr>
<tr>
<td>No corner</td>
<td>0.030</td>
<td>0.031</td>
<td>0.013</td>
<td>0.032</td>
<td>0.012</td>
<td>0.031</td>
</tr>
<tr>
<td>Entrance</td>
<td>(2.004)</td>
<td>(2.024)</td>
<td>(0.896)</td>
<td>(2.063)</td>
<td>(0.846)</td>
<td>(1.999)</td>
</tr>
<tr>
<td>Space</td>
<td>0.020</td>
<td>0.020</td>
<td>0.013</td>
<td>0.022</td>
<td>0.013</td>
<td>0.023</td>
</tr>
<tr>
<td></td>
<td>(1.979)</td>
<td>(2.012)</td>
<td>(0.759)</td>
<td>(0.926)</td>
<td>(0.752)</td>
<td>(0.954)</td>
</tr>
<tr>
<td>No shared</td>
<td>0.017</td>
<td>0.017</td>
<td>0.033</td>
<td>0.002</td>
<td>0.034</td>
<td>0.001</td>
</tr>
<tr>
<td>Building</td>
<td>(1.284)</td>
<td>(1.288)</td>
<td>(2.061)</td>
<td>(0.120)</td>
<td>(2.090)</td>
<td>(0.078)</td>
</tr>
<tr>
<td>Large sign</td>
<td>0.022</td>
<td>0.022</td>
<td>0.568</td>
<td>0.562</td>
<td>0.577</td>
<td>0.553</td>
</tr>
<tr>
<td>Constant</td>
<td>0.546</td>
<td>0.545</td>
<td>0.870</td>
<td>0.789</td>
<td>0.870</td>
<td>0.789</td>
</tr>
<tr>
<td></td>
<td>(18.731)</td>
<td>(18.804)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>F-test, p-value</td>
<td>0.001</td>
<td>0.001</td>
<td>0.024</td>
<td>0.164</td>
<td>0.022</td>
<td>0.163</td>
</tr>
<tr>
<td>R²</td>
<td>0.834</td>
<td>0.834</td>
<td>Note</td>
<td>High pop.</td>
<td>Low pop.</td>
<td>Low pop.</td>
</tr>
<tr>
<td>Sample size</td>
<td>536</td>
<td>536</td>
<td>268</td>
<td>268</td>
<td>268</td>
<td>268</td>
</tr>
</tbody>
</table>

Notes: The dependant variable is the TANF ratio, computed in the same way as above. All specification include unreported controls for state and poverty rate. Columns (1) and (2) interact busy street with population and also have unreported controls for population, as well as for the black share of the population in Column (2). Columns (3) to (6) divide the sample into high- and low-population welfare offices. Columns (5), and (6) also control for county population. The absolute values of the associated heteroscedasticity-robust t-statistics are given within parentheses. The F-test refers to a test of joint significance of all the building characteristics, whether interacted or no. Heteroscedasticity-robust t-statistics are given within parentheses. The Bonferroni-adjusted t-tests' critical values for statistical significance are approximately 2.46 (ten per cent), 2.70 (five per cent), and 3.20 (one per cent), depending very slightly on sample sizes. This adjustment corrects for testing seven building characteristics.

It may be interesting to compare coefficients across states to see that one or two states are not pulling most of the weight, even though the sample sizes become rather small. Doing so yields the expected finding that f-statistics generally go down significantly with no individual variable being statistically significant at even the ten per cent level in all states, and that the signs and magnitudes of the building coefficients, while still generally supportive of the hypothesis, come to be more variable. With smaller samples, such variations in point estimates are also a natural consequences of the characteristics being all dummies. Estimates are reported in Table IV.

Neighbourhood income measures average income in the area immediately surrounding a welfare office and is available for the Georgia subsample (N=148). Although the sample size is rather small, controlling for it has absolutely no effect with a barely positive coefficient and a p-value of 0.936 (controlling also for county poverty
rate). This is consistent with the insignificant findings reported above on welfare offices’ proximity to the town centre or fraction of service area living in the town of the welfare office, and further underscores the efforts by Human Services departments at locating in a uniform manner with respect to factors such as easy access and cleanliness.

Most economic norm theory predicts that the reluctance to be seen to apply for welfare benefits is reduced and perhaps eventually overturned when sufficiently few people are self-reliant. One way of estimating this effect is to look at how building effects concentrate at different levels of income by cutting the sample into two halves at the median county income ($39,872). Table V presents these alternative results, where odd columns display the sample in which income is above the median, and even columns the half in which income is below.

Although the signs remain mostly positive in the relatively low-income sample, the F-tests on all building characteristics invariably have p-values close to 0.9, whereas in odd columns, the p-values on the F-tests never go any higher than 0.003. The relatively high-income subsample also has uniformly greater coefficients (except, again, busy street), indicating that the norm of self-reliance does indeed weaken as poverty increases. The (weak) evidence provided by these analyses is that the repulsive effects of visibility-enhancing building characteristics seem to diminish as the rate of poverty goes up, and consequently that some small fraction of welfare offices in high-poverty areas may stand to gain some clients by making entry more conspicuous.
As an alternative to individual visibility-enhancing building characteristics, Table VI (a) presents estimates using a visibility index, ranging from zero to six (dropping the population-sensitive variable busy street), in which the building dummies are simply summed (mean 3.04; median 3). Out of 536 observations, 473 are between two and four on this index, and 254 are threes. Also, there are only three noughts and three sixes (excluding them would yield nearly the same coefficient).

The visibility-indexed results are similar to the ones based on individual building characteristics. In columns (7) and (8), the same analyses on high- and low-income subsamples are performed as in Table V. Using the visibility index in lieu of individual building variables strengthens the impression that being seen to go to the welfare office is worse in affluent areas than less well-to-do places. Column (6) complements the analysis by subsamples by interacting the visibility index with the poverty rate. The county rate of poverty at which visibility-enhancing building characteristics go from repelling to attracting applicants is predicted to be around forty-two per cent, which is substantially higher than the 35.8-per-cent poverty rate found in the poorest sampled county, but it should be noted that the t-statistic on the interaction term is very low.

Table VI (c) performs a principal component analysis of the visibility-enhancing variables. The first four principal components (PC 1 to PC 4) account for almost three quarters of the variance, and the first five for almost ninety per cent. The proportion of the variance accounted for by the first principal component is less than twice that of the sixth component. Certain natural tendencies can be seen in the analysis. Space and the

<table>
<thead>
<tr>
<th>Table V: Estimates on Interaction Effects with Income</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
</tr>
<tr>
<td>Busy street</td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td>Windows</td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td>Elevated entrance</td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td>No corner</td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td>Space</td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td>No shared building</td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td>Large sign</td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td>Poverty rate</td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td>County population</td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td>F-test, p-value</td>
</tr>
<tr>
<td>R²</td>
</tr>
<tr>
<td>Sample size</td>
</tr>
</tbody>
</table>

Notes: The dependant variable is the TANF ratio, computed in the same way as before. All regressions include state controls and a constant, whose coefficients are not reported but are similar to the ones obtained in Table II. The sample is divided along the median county household income ($39,872), so that odd columns include only counties with a median household income higher than this, and even columns include the other, lower-income, half. The absolute values of the associated heteroscedasticity-robust t-statistics are given within parentheses. The Bonferroni-adjusted t-tests’ critical values for statistical significance are approximately 2.47 (ten per cent), 2.71 (five per cent), and 3.22 (one per cent), depending very slightly on sample sizes. This adjustment corrects for testing seven building characteristics.
absence of a corner entrance seem to be the main drivers of the variance accounted for by the first principal component, where Space is also negatively correlated with Windows: it is harder to look inside from a greater distance. Not sharing a building with non-welfare office outfits also seems to be correlated with having a large sign.

Table VI (a): Results Using a Visibility Index in Lieu of Individual Building Characteristics

<table>
<thead>
<tr>
<th></th>
<th>(1)</th>
<th>(2)</th>
<th>(3)</th>
<th>(4)</th>
<th>(5)</th>
<th>(6)</th>
<th>(7)</th>
<th>(8)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Visibility</td>
<td>0.016</td>
<td>0.023</td>
<td>0.020</td>
<td>0.018</td>
<td>0.022</td>
<td>0.027</td>
<td>0.027</td>
<td>0.008</td>
</tr>
<tr>
<td></td>
<td>(1.279)</td>
<td>(2.087)</td>
<td>(3.768)</td>
<td>(3.480)</td>
<td>(4.180)</td>
<td>(1.890)</td>
<td>(4.350)</td>
<td>(0.977)</td>
</tr>
<tr>
<td>Visibility*Pov.  Rate</td>
<td>-0.054</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td>(0.497)</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Distance fr.</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-0.002</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Centre</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Fraction in</td>
<td>-0.005</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Town</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>State controls</td>
<td>No</td>
<td>No</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>Poverty rate</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td>(9.311)</td>
<td>(4.275)</td>
<td>(4.184)</td>
<td>-</td>
<td>(2.014)</td>
<td>(0.277)</td>
<td>(3.153)</td>
<td></td>
</tr>
<tr>
<td>R²</td>
<td>0.003</td>
<td>0.143</td>
<td>0.830</td>
<td>0.823</td>
<td>0.823</td>
<td>0.830</td>
<td>0.885</td>
<td>0.763</td>
</tr>
<tr>
<td>Sample size</td>
<td>536</td>
<td>536</td>
<td>536</td>
<td>520</td>
<td>536</td>
<td>536</td>
<td>268</td>
<td>268</td>
</tr>
<tr>
<td>Note:</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Table VI (b): The Effect of Increased Online Application Possibilities

<table>
<thead>
<tr>
<th>Dep. variable</th>
<th>After reform</th>
<th>After reform</th>
<th>Before reform</th>
<th>Before reform</th>
<th>Difference</th>
<th>Difference</th>
</tr>
</thead>
<tbody>
<tr>
<td>Visibility</td>
<td>0.008</td>
<td>0.003</td>
<td>0.028</td>
<td>0.025</td>
<td>-0.018</td>
<td>-0.017</td>
</tr>
<tr>
<td></td>
<td>(0.636)</td>
<td>(0.222)</td>
<td>(2.060)</td>
<td>(1.835)</td>
<td>(1.441)</td>
<td>(1.367)</td>
</tr>
<tr>
<td>Poverty rate</td>
<td>-</td>
<td>1.534</td>
<td>-</td>
<td>0.715</td>
<td>-</td>
<td>-0.422 (0.546)</td>
</tr>
<tr>
<td></td>
<td>-</td>
<td>(3.430)</td>
<td>-</td>
<td>(1.750)</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>R²</td>
<td>0.620</td>
<td>0.645</td>
<td>0.575</td>
<td>0.581</td>
<td>0.028</td>
<td>0.030</td>
</tr>
</tbody>
</table>

Table VI (c): Principal Component Analysis of Building Variables in the Full Sample

<table>
<thead>
<tr>
<th>Component</th>
<th>Eigenvalue</th>
<th>Proportion</th>
<th>Variable</th>
<th>PC 1</th>
<th>PC 2</th>
<th>PC 3</th>
<th>PC 4</th>
<th>PC 5</th>
<th>PC 6</th>
</tr>
</thead>
<tbody>
<tr>
<td>Component 1</td>
<td>1.2978</td>
<td>0.2163</td>
<td>Windows</td>
<td>-0.473</td>
<td>0.023</td>
<td>0.468</td>
<td>0.531</td>
<td>0.404</td>
<td>0.334</td>
</tr>
<tr>
<td>Component 2</td>
<td>1.1599</td>
<td>0.4096</td>
<td>Elev. ent.</td>
<td>0.184</td>
<td>0.217</td>
<td>0.764</td>
<td>-0.547</td>
<td>0.141</td>
<td>-0.131</td>
</tr>
<tr>
<td>Component 3</td>
<td>1.0089</td>
<td>0.5578</td>
<td>No corner</td>
<td>0.513</td>
<td>0.020</td>
<td>0.292</td>
<td>0.640</td>
<td>-0.212</td>
<td>-0.444</td>
</tr>
<tr>
<td>Component 4</td>
<td>0.9488</td>
<td>0.7359</td>
<td>Space</td>
<td>0.672</td>
<td>0.072</td>
<td>-0.069</td>
<td>0.048</td>
<td>0.222</td>
<td>0.698</td>
</tr>
<tr>
<td>Component 5</td>
<td>0.8462</td>
<td>0.8769</td>
<td>No sh. bldg</td>
<td>0.039</td>
<td>0.685</td>
<td>-0.323</td>
<td>0.057</td>
<td>0.564</td>
<td>-0.323</td>
</tr>
<tr>
<td>Component 6</td>
<td>0.7383</td>
<td>1.0000</td>
<td>Large sign</td>
<td>-0.162</td>
<td>0.691</td>
<td>0.060</td>
<td>0.070</td>
<td>-0.636</td>
<td>0.288</td>
</tr>
</tbody>
</table>

Notes: The dependant variable in panel (a) is the TANF ratio, computed in the same way as above. All regressions include an unreported constant. The variable Fraction in town is the fraction of the population in a service area living in the town of the welfare office. Column (7) shows results from the higher-income half of the sample and Column (8) results from the lower-income half. In panel (b), the visibility variable excludes busy street and the regressions are performed on 162 observations from North Carolina and Oklahoma. The two rightmost Difference-columns are the ‘after’ ratios minus the ‘before’ ratios. The absolute values of the associated heteroscedasticity-robust t-statistics are given within parentheses. The t-tests’ critical values for statistical significance are approximately 1.65 (ten per cent), 1.96 (five per cent), and 2.60 (one per cent), depending very slightly on sample sizes. Panel (c) shows a principal component analysis. “PC N” refers to the nth principal component and “Proportion” is the cumulative proportion of the variation explained by a component.

Table VI (b) examines the possibility to apply for food stamps online in North Carolina and Oklahoma. Between 2010 and 2013, North Carolinians could gradually look at information about welfare programmes as well as completing applications for certain benefits, most notably food stamps, through an online facility (ePASS). Oklahomans could do the same, though not gradually, also during this period through OKDHS Live. Only more recent Oklahoman data on online applications are available. They indicate
that roughly 50-65 per cent of applications are made online. Unfortunately, these data cannot be broken down by welfare office.

While food stamps are separate from TANF, failing to be approved for the former (a much bigger programme) plausibly gives the applicant an indication of his likely success were he to apply for the latter. Consequently, Table VI (b) shows before-and-after comparisons of the approval ratio for TANF payments in North Carolina (NC) and Oklahoma (OK), where ‘before’ means that the ratio is from 2011 (OK) or 2010 (NC), and ‘after’ means 2014 (OK), or July 2013-June 2014, with the exception of five counties instead sampled in 2013 (Jan.-Dec.) due to lack of available data for later months (NC).

The results show no indication that the effects from visibility should disappear. The visibility variable is essentially zero after the introduction of ePASS or OKDHS Live and uniformly positive before. Subtracting the 2010 ratios from the ’13-’14 ones, as in the two rightmost columns of panel (b), reveals that the approval rate increases most strongly in counties whose welfare offices have fewer visibility-enhancing building characteristics. The buildings remained the same while welfare application procedures changed, so this result again indicates that visibility matters. Regressions with individual building characteristics are unreported but similar in essence; the variables are mainly positive even in 2013-14, but less so and with lesser statistical significance.
IV – DISCUSSION

There are more states than the sampled ones which require in-person visits to a welfare office when applying for TANF, but the six sampled states are the only ones who shared their data. The nature of the results is such that one may suspect stronger statistical associations with a larger sample. While very much consistent with established theories of social influences on individuals’ behaviour, these results are first and foremost correlations. Unfortunately, many good methods that strengthen identification are not available. Such methods might include looking at application behaviour in areas straddling service-area borders to determine whether similar neighbours who have to frequent different-looking welfare offices apply with different degrees of conservatism, or finding suitable instrumental variables. With data on application days, unusually foggy periods may exogenously impact visibility.

One potential source of bias is that TANF applicants may choose to move where welfare offices allow for least conspicuous entries. If those who are most certain of having their applications accepted are also most inclined to move, this force would lower the approval rate in sending service areas with many visibility-enhancing building characteristics. Since visibility is hypothesized to be associated with a high approval rate, this force would make welfare shame harder to detect in the present paper and so the results are likely an underestimate.

Another possible source of bias is that, contrary to official guidelines, certain service areas may be less well funded and thereby see an automatic decrease in the approval rate. While the funding a welfare office receives is implausibly reduced simply because it has few visibility-enhancing building characteristics, any chance tendency in that direction will cause the results to be overstated. This paper’s utilization of the move towards online welfare services withstands this potential source of bias to the extent that Internet-given anonymity eliminates the need for persons to physically appear at their welfare offices. If online services continue to expand, perhaps to also include TANF applications, they will eventually yield greater sample sizes than that of the present paper.

The present study goes well with a great deal of the established literature on peer effects. For instance, under the complementarities approach to peer effects (e.g. Becker and Murphy, 2000), one may think of two ways of obtaining income, by social assistance, \( w \), or by work, \( e \). To incorporate a social aspect to this setting, surviving on welfare can be considered a complement to a non-choice variables \( d \), capturing average reliance on social assistance among one’s peers (as in Lindbeck et al., 1999), and \( v \), capturing the visibility of making a welfare application. In the present analysis, visibility-enhancing building characteristics essentially shift the utility of relying on welfare payments. For low rates of welfare seeking (i.e. under a strong norm of self-reliance), the shift is to lower utility when inconspicuous entry is hard, and to higher utility when easy. The estimates from Section III suggest that the effect of visibility may be reversed in areas of particularly high poverty.

If an individual gains utility in this manner, his utility function is of the form

\[
U = U(w, e; d; v),
\]

where both earned income \( e \) and welfare \( w \) yield utility, and where \( \frac{\partial^2 U}{\partial w \partial d} > 0 \) and \( \frac{\partial^2 U}{\partial w \partial d}|_{d<d^*} < 0 \), so that increased reliance on welfare makes it better for the individual to be on welfare, and visibility reduces welfare for persons making a welfare application whose peers rely on welfare to a degree \( d \) below the critical value \( d^* \) at which visibility possibly becomes utility-enhancing: \( \frac{\partial^2 U}{\partial w \partial v}|_{d>d^*} > 0 \). Letting a finite time endowment act
as a budget constraint with one good as numéraire and the other priced at $p$ and given the usual assumptions about quasi-concavity of the utility function, it can be shown that the ‘demand’ function for living on welfare is negative when price-adjusted marginal utility of living on one’s own efforts declines more rapidly in average reliance on the dole than does the marginal utility of living on social assistance: \[
\frac{\partial w}{\partial d} = -\frac{pU_{cd} - U_{wd}}{\text{denominator} > 0}.
\]

Similarly, the ‘demand’ function for living on welfare is affected by changes in visibility by \[
\frac{\partial w}{\partial v} = \frac{U_{vw} - pU_{ev}}{\text{denominator} > 0}.\]

When welfare reliance is not a norm ($d < d^*$), $U_{vw} < 0$ since higher visibility reduces whatever utility that comes from the prospect of receiving a transfer. Provided that one’s decision to earn income, $e$, is either made more appealing (which is equivalent to stigma being a normal bad) or is not affected by greater visibility of seeking a transfer, $v$, $U_{ev}$ is non-negative and visibility always reduces an individual’s propensity to apply for welfare given $d < d^*$.

An exogenous increase in visibility lowers demand for welfare by more the more it raises the price-adjusted marginal utility of earning income relative to that of living on welfare. *Ceteris paribus*, this effect is greater for applicants whose chances of approval are relatively low, than it is for applicants with high chances of approval. The former’s substitution of time spent working for time spent seeking welfare (talking to caseworkers, meeting requirements for payments, etc.) proves their relative self-reliance and thereby their lesser need for assistance. Consequently, a rise in visibility will be accompanied by an increase in the acceptance-to-applications ratio.

It is clear from the hypothesis that entry decisions depend on the expected reward, in the form of social assistance, compared to the cost of visibility. Since the expected reward is the probability of being approved multiplied by the sum awarded, holding the former constant while varying the latter should reveal that ‘small’ payments are no more commonly received when applicants go to a high-visibility welfare office compared to a low-visibility one. That is, the low-payment tail of the distribution is cut by visibility-enhancing building characteristics. This testable implication is not explored in the present article due to data limitations, but it would be a feasible empirical project.

Another implication is that the same essential phenomenon of social opprobrium and aversion that varies with visibility may be expected in highly disparate settings. For example, many individuals feel pressure to maintain a healthy image, so self-service check-outs in supermarkets may be more popular among persons purchasing items which are considered low-status along this dimension. Another example concerning health is that certain diseases are embarrassing. Going to the proctologist’s office or to a centre testing for social diseases is therefore apt to reduce one’s status and people may be more likely to avoid such visits if the waiting rooms have glass walls, say. Whether this is good or bad depends on the prudence with which such decisions are otherwise made. ‘Free’ clinics may also pay less heed to such social concerns than for-profit hospitals who gain by getting more visits.
V – CONCLUSION

Related to the present topic, it is not uncommon for famous individuals from poor backgrounds to tell stories from their childhood of how they would alight at bus stops adjacent to their nearest ones because they felt ashamed of where they lived. Although the evidence in the present article comes from a different source of shame, it suggests that the picture painted by anecdotes like these is well-founded.

The present analysis also suggests analogous treatments of yet other activities which indicate, or are associated with, social stigma. Several avenues for further investigations suggest themselves. For instance, might the same effects be found in testing centres for social diseases? Do such clinics receive more clients if they are part of a general-practice surgery? Have self-service check-outs at supermarkets contributed to purchases of less salutary victuals (since social diseases and unhealthful eating are commonly regarded as ‘shameful’)? Etc. In the instances mentioned, visibility can plainly vary a great deal, so these issues could have significant impacts on public health. The present article, too, is policy relevant in its relation to welfare ordeals.

Although the present analysis cannot randomize building characteristics, utilization of control variables, background information, and application reforms all yield results consistent with social influences on behaviour; many individuals do appear to be sufficiently sensitive to welfare stigma to affect the approval rate for Temporary Assistance for Needy Families. Seemingly insignificant architectural traits such as corner entrances, a few steps leading up to the door, or an entrance located farther away from the street, appear to promise small but positive savings on the time social workers spend reviewing applications.

Also noteworthy is the result that a higher poverty rate is in line with weakly ‘normalizing’ welfare applications. These results are qualitatively less strong and statistically less significant than are the main ones. However, the overall thrust of the results suggests that the level of poverty is a complicating factor in individuals’ responses to welfare stigma. At any rate, buildings seem to be important also for reasons altogether different from giving shelter.
BIBLIOGRAPHY


Department of Human Services, Oklahoma, New and denied TANF applications, available from the DHS Oklahoma website: http://www.okdhs.org/, where the welfare offices’ addresses can also be found.

Departments of Human Services (or Human Resources) in Alabama, Georgia and Oregon provide data on welfare applications on request and in the case of Georgia for a fee. The addresses of their welfare offices are available from these departments’ websites: http://dhr.alabama.gov/, http://dhs.georgia.gov/, and http://www.oregon.gov/dhs/.

Department of Social Services, Missouri, Applications data from 2013 available from the Missouri Family Support Division Annual Data Report, online at: http://dss.mo.gov/re/fsdar.htm.


PUZZLING EVIDENCE ON VOTER TURNOUT

Mats Ekman

In this empirical analysis of voting patterns in five countries on days when one or more national referenda were held, voter turnout appears to decline in the number of concurrent referenda, in contrast to standard theories’ predictions and regardless of method used to hold constant the quality of the referenda. Multiple concurrent referenda imply “quantity discounts” as one may vote on more ballots in one visit to the polling station. They should also draw more voters due to the wider range of interests attracted when more issues are up for vote. Yet, none of this seems to happen in the data. More recent developments, such as rule-utilitarian and information-based theories of voting, fare similarly poorly in light of the evidence presented in this article; a social theory of voting does better. (JEL code: D72)

---

8 This chapter is published in *Rationality and Society*, Vol. 29, Issue 4 (Nov., 2017), pp. 449-470. Special thanks to Tuomas Nurminen, Rune Stenbacka, Janne Tukiainen, and two anonymous referees for helpful discussions or suggestions. The usual disclaimer applies. The support of the Society of Swedish Literature in Finland is gratefully acknowledged.
I - INTRODUCTION

Look at Figure I. It displays a naïve regression of voter turnout on the number of national referenda held on the same day, a total of 383 instances of direct democracy, of which 279 are from Switzerland, 36 from New Zealand, 25 from Ireland, 22 from Australia, and 21 from Italy. These votes, were taken between 1889 and 2014, 199 of them are single-ballot days and 90 of the remaining 184 instances are days on which two referenda were held simultaneously. (The highest number of concurrent referenda in the sample is twelve, from Italy.) The figure indicates that, contrary to standard theoretical predictions, holding more than one referendum on the same day fails to attract higher turnout. While only a naïve regression, the gist of its message withstands the addition of standard control variables and ways of holding constant the quality of the referenda, and is robust across countries.

Voter turnout is an important issue for a variety of reasons. The tiny chance of casting a pivotal ballot has long attracted interest in why rational individuals vote at all, and who decides to vote plainly impacts the outcome of a referendum. Standard theories of voter turnout utilize a “calculus of voting” in which the individual compares the costs of voting to the expected benefits of doing so. A vast literature has resulted from this approach; if more individuals perceive a vote to be more important, they should be more likely to vote; better access to the polls should similarly increase turnout by lowering voting costs, etc. (Dhillon and Peralta, 2002; Grossman and Helpman, 2001; Farber, 2010; McMurray, 2013; Lyytikäinen and Tukiainen, 2013; Andersen et al, 2014; Herrera and Morelli, 2014).

These theories also predict that an individual is more likely to vote when he can cast multiple ballots at a time. For instance, a multitude of concurrent referenda is apt to attract a wider span of interests than is a stand-alone referendum, and should therefore get higher turnout. In addition, the same stroll to the polling station (or post box in the case of postal voting) influences more issues on days of multiple concurrent referenda. Section IV below adds a more detailed theoretical treatment. On its face, the relationship depicted in Figure I is akin to quantity discounts being associated with the purchase of fewer products in the market.

In spite of all the interest in voter turnout, estimates on the effects of concurrent votes tend to come from elections, not referenda, and turnout in elections is typically found to rise in the presence of more ballots (see, e.g., Fauvelle-Aymar and François, 2015, who
utilize a peculiarity in French local elections that lets some voters in a constituency vote in a departmental and a regional election; or see Geys, 2006, for a review).

Focusing on referenda therefore adds a novel dimension to the empirics of voter turnout. Referenda may treat issues like defence and helmet laws whereas all elections have a strong partisan element. Elections are therefore less likely than referenda to change the composition of groups with a high stake in the result. If many people vote mainly to please their peers, who (because of the high stakes) punish non-voters at a cost, the contrasting empirics on turnout between referenda and elections may be explained by group-level free-riding as the probability that an abstainer also belongs to a different group increases the more concurrent referenda there are. If another group can carry out the costly punishment, nobody punishes anyone and abstainers escape scot-free. Section IV below will discuss this (tentative) explanation in greater detail.

The rest of this article proceeds as follows: data and institutional background are treated in the next section, whereupon Section III presents the main results, which corroborate the testimony of the introductory regression. Section IV discusses how different explanations of turnout, established ones and variations on them, fare in the face of these results, and elaborates on the group-based free-riding explanation mentioned above as a potential way of making sense of the empirical findings. Section V concludes.
<table>
<thead>
<tr>
<th>Turnout (full sample)</th>
<th>Mean</th>
<th>Median</th>
<th>St. dev.</th>
<th>Sample size</th>
<th>Year range</th>
</tr>
</thead>
<tbody>
<tr>
<td>Turnout (single referendum)</td>
<td>59.803</td>
<td>57.060</td>
<td>18.127</td>
<td>199</td>
<td>1889-2012</td>
</tr>
<tr>
<td>Turnout (multiple referenda)</td>
<td>49.883</td>
<td>46.190</td>
<td>15.526</td>
<td>184</td>
<td>1891-2014</td>
</tr>
<tr>
<td>Turnout (Switzerland)</td>
<td>49.531</td>
<td>47.640</td>
<td>12.330</td>
<td>279</td>
<td>1889-2014</td>
</tr>
<tr>
<td>Single-referendum turnout (Switzerland)</td>
<td>53.481</td>
<td>51.790</td>
<td>13.949</td>
<td>131</td>
<td>1889-2012</td>
</tr>
<tr>
<td>Multiple-referenda turnout (Switzerland)</td>
<td>46.034</td>
<td>45.245</td>
<td>9.445</td>
<td>148</td>
<td>1891-2014</td>
</tr>
<tr>
<td>Turnout (Australia)</td>
<td>83.987</td>
<td>92.835</td>
<td>15.263</td>
<td>22</td>
<td>1900-1999</td>
</tr>
<tr>
<td>Single-referendum turnout (Australia)</td>
<td>81.490</td>
<td>88.155</td>
<td>17.846</td>
<td>8</td>
<td>1900-1951</td>
</tr>
<tr>
<td>Multiple-referenda turnout (Australia)</td>
<td>85.414</td>
<td>92.835</td>
<td>14.099</td>
<td>14</td>
<td>1910-1999</td>
</tr>
<tr>
<td>Single-referendum turnout (Ireland)</td>
<td>54.063</td>
<td>53.400</td>
<td>10.659</td>
<td>18</td>
<td>1937-2012</td>
</tr>
<tr>
<td>Multiple-referenda turnout (Ireland)</td>
<td>51.461</td>
<td>55.960</td>
<td>14.879</td>
<td>7</td>
<td>1968-2011</td>
</tr>
<tr>
<td>Turnout (Italy)</td>
<td>55.476</td>
<td>54.830</td>
<td>22.581</td>
<td>21</td>
<td>1946-1993</td>
</tr>
<tr>
<td>Single-referendum turnout (Italy)</td>
<td>63.113</td>
<td>62.500</td>
<td>21.763</td>
<td>9</td>
<td>1946-2007</td>
</tr>
<tr>
<td>Multiple-referenda turnout (Italy)</td>
<td>49.748</td>
<td>49.095</td>
<td>22.341</td>
<td>12</td>
<td>1978-2011</td>
</tr>
<tr>
<td>Turnout (New Zealand)</td>
<td>80.945</td>
<td>84.235</td>
<td>13.941</td>
<td>36</td>
<td>1896-2011</td>
</tr>
<tr>
<td>Single-referendum turnout (New Zealand)</td>
<td>81.867</td>
<td>84.660</td>
<td>13.764</td>
<td>33</td>
<td>1896-2011</td>
</tr>
<tr>
<td>Multiple-referenda turnout (New Zealand)</td>
<td>70.800</td>
<td>71.200</td>
<td>14.204</td>
<td>3</td>
<td>1949-1999</td>
</tr>
</tbody>
</table>

Additionally, the number of twin, triplet, etc., referenda in each country are the following (twins; triplets; ...):
Switzerland (71; 40; 18; 13; 4; 1; 0; 1); New Zealand (3); Ireland (5; 2); Australia (9; 1; 2; 1; 1); Italy (2; 2; 2; 0; 2; 1; 0; 0; 1).
II – DATA AND INSTITUTIONAL BACKGROUND

Referenda are common in political decision-making on a wide variety of issues across the world, but multiple concurrent referenda on the national stage occur less frequently. Table I provides summary statistics on the five countries studied in this paper: Australia, Ireland, Italy, New Zealand, and Switzerland. There are very few instances of multiple concurrent national referenda other than these. Average turnout in the sample is almost unfailingly lower for multiple concurrent referenda than for single ones, except in Ireland, whose median turnout figure deviates slightly (Australia differs on both measures only because ten out of the 14 observations of multiple referenda are from after it introduced compulsory voting in 1924; the same share for its stand-alone referenda is four out of eight).

Turnout is measured as the votes cast divided by the total number of eligible voters. On days of multiple concurrent referenda, turnout may vary across ballots, in which case the highest turnout is chosen to represent all the referenda held on that day (however, the variation is never greater than a fraction of a percentage point and usually zero). The voting day is always a Sunday in Switzerland, but varies otherwise. Day-of-the-week dummies in the regressions below are never reported; adding them does not qualitatively change the results, but the smaller-sampled regressions lose some statistical significance. The same goes for month of the year; results for which are also left unreported.

Occasionally, parliamentary elections accompany referenda in the sampled countries. These often – but not always – raise turnout, although not statistically significantly. In Italy, there has been one instance of a referendum held jointly with a parliamentary election (in 1946); in Ireland it has happened twice; in New Zealand it is the norm with only eight exceptions; there are seven cases in Australia and five in Switzerland.

In addition, and more importantly, the citizens’ (voters and abstainers) perceived importance of referenda has been surveyed in Switzerland since the 1980’s and offer a way of holding otherwise potentially disparate referenda constant. The issues with which the subsample of Swiss referenda is concerned are also available and offer supporting though weaker control variables.

The variable Victory Margin is the percentage-point difference between the winning and losing alternatives. In observations of multiple concurrent referenda, the narrowest victory margin is used. In nine cases, the margin of victory is set to zero due to referenda rules stipulating qualified majorities of various kinds; for instance, if a proposal received 51 per cent of the popular vote, but failed to receive majorities in a most of states or cantons. In other cases, turnout did not reach a threshold necessary for the referendum to be valid. Here, the victory margin is the shortfall in required turnout and yes-votes (if needed).

As indicated above, voting is mandatory in some jurisdictions. In Australia, abstainers have been fined since 1924. In Italy, abstention from voting was punished by publication of one’s name until 1993 (Malkopoulou, 2014). Swiss non-voters were fined from 1904 until 1974 (López-Pintor and Gratschew, 2004, p. 74) and one of its 26 cantons, Schaffhausen, still levies a six-franc fine for abstention, although reports indicate that any fool will have an easy time avoiding punishment. It is, however, doubtful whether

---

9 Almost all the data come from the Centre for Research on Direct Democracy (URL: http://www.czd.ch/), which collects information on turnout, such as a brief description of the proposals up for vote, referendum dates, and victory margins. Complementary data (post-referendum surveys of respondents’ perceived importance of the recent referendum or referenda) come from the "VOXIT" surveys available from the Swiss Foundation for Research in Social Sciences (http://fors-nestar.unil.ch/webview/index.jsp). With the occasional exception of 25 observations from New Zealand (which miss values for margins of victory), the data include every referendum held since the sample's starting point, which coincides with the first referendum on which complete data are available.

10 See the article ‘Democratic? The Canton where Voting is Compulsory’ (April 4th, 2014) by Isobel Leybold-
the practice of compulsory voting ever significantly affected Swiss turnout rates. Partly for this reason, Section III spends some time comparing the effect of compulsory voting on turnout rates in Switzerland to Australia and Italy. The conclusion is that compulsory voting significantly increases turnout in Italy and Australia, but has no effect on Switzerland, indicating that Swiss estimates should assume no compulsory voting.

Table II: OLS Results for Various Specifications with Turnout as the Dependant Variable

<table>
<thead>
<tr>
<th></th>
<th>N=358</th>
<th>N=358</th>
<th>N=279</th>
<th>N=279</th>
<th>N=358</th>
<th>N=358</th>
<th>N=279</th>
<th>N=279</th>
</tr>
</thead>
<tbody>
<tr>
<td>Year of Vote</td>
<td>-0.159</td>
<td>-0.161</td>
<td>-0.175</td>
<td>-0.176</td>
<td>-0.115</td>
<td>-0.113</td>
<td>-0.168</td>
<td>-0.166</td>
</tr>
<tr>
<td>(7.264)</td>
<td>(7.460)</td>
<td>(7.270)</td>
<td>(7.345)</td>
<td>(3.700)</td>
<td>(3.581)</td>
<td>(7.188)</td>
<td>(7.103)</td>
<td></td>
</tr>
<tr>
<td>Victory Margin (in percentage points)</td>
<td>-0.228</td>
<td>-0.241</td>
<td>-0.203</td>
<td>-0.203</td>
<td>-0.260</td>
<td>-0.269</td>
<td>-0.203</td>
<td>-0.205</td>
</tr>
<tr>
<td>(2.976)</td>
<td>(3.187)</td>
<td>(2.128)</td>
<td>(2.154)</td>
<td>(3.189)</td>
<td>(3.354)</td>
<td>(2.149)</td>
<td>(2.193)</td>
<td></td>
</tr>
<tr>
<td>Australia</td>
<td>7.628</td>
<td>8.302</td>
<td>28.507</td>
<td>28.705</td>
<td>–</td>
<td>–</td>
<td></td>
<td></td>
</tr>
<tr>
<td>(1.900)</td>
<td>(2.044)</td>
<td>–</td>
<td>(4.700)</td>
<td>(4.745)</td>
<td>–</td>
<td>–</td>
<td></td>
<td></td>
</tr>
<tr>
<td>New Zealand</td>
<td>22.404</td>
<td>22.888</td>
<td>25.701</td>
<td>25.629</td>
<td>–</td>
<td>–</td>
<td></td>
<td></td>
</tr>
<tr>
<td>(7.370)</td>
<td>(7.601)</td>
<td>–</td>
<td>(8.139)</td>
<td>(8.383)</td>
<td>–</td>
<td>–</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ireland</td>
<td>7.952</td>
<td>8.129</td>
<td>10.994</td>
<td>10.772</td>
<td>–</td>
<td>–</td>
<td></td>
<td></td>
</tr>
<tr>
<td>(3.736)</td>
<td>(3.827)</td>
<td>–</td>
<td>(4.509)</td>
<td>(4.497)</td>
<td>–</td>
<td>–</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Italy</td>
<td>-4.716</td>
<td>-3.397</td>
<td>10.939</td>
<td>11.088</td>
<td>–</td>
<td>–</td>
<td></td>
<td></td>
</tr>
<tr>
<td>(1.557)</td>
<td>(1.293)</td>
<td>–</td>
<td>(1.539)</td>
<td>(1.684)</td>
<td>–</td>
<td>–</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Compulsory Voting</td>
<td>34.072</td>
<td>33.261</td>
<td>7.637</td>
<td>7.525</td>
<td>1.641</td>
<td>1.678</td>
<td></td>
<td></td>
</tr>
<tr>
<td>(9.229)</td>
<td>(8.875)</td>
<td>–</td>
<td>(2.896)</td>
<td>(2.930)</td>
<td>(0.843)</td>
<td>(0.903)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>(1.393)</td>
<td>(1.282)</td>
<td>(0.809)</td>
<td>(0.879)</td>
<td>(1.502)</td>
<td>(1.477)</td>
<td>(0.697)</td>
<td>(0.741)</td>
<td></td>
</tr>
<tr>
<td>Parliamentary Election</td>
<td>-3.116</td>
<td>-2.580</td>
<td>-1.138</td>
<td>-2.159</td>
<td>–</td>
<td>–</td>
<td></td>
<td></td>
</tr>
<tr>
<td>(3.171)</td>
<td>(1.733)</td>
<td>–</td>
<td>(1.089)</td>
<td>(1.458)</td>
<td>–</td>
<td>–</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Number of Referenda</td>
<td>0.339</td>
<td>0.285</td>
<td>0.113</td>
<td>0.250</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(3.543)</td>
<td>(1.621)</td>
<td>–</td>
<td>(0.991)</td>
<td>(1.445)</td>
<td>–</td>
<td>–</td>
<td></td>
<td></td>
</tr>
<tr>
<td>(2.322)</td>
<td>(1.666)</td>
<td>–</td>
<td>(1.398)</td>
<td>(1.427)</td>
<td>–</td>
<td>–</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Two or More Referenda (dummy)</td>
<td>–</td>
<td>370.627</td>
<td>399.749</td>
<td>275.719</td>
<td>270.816</td>
<td>383.100</td>
<td>378.920</td>
<td></td>
</tr>
<tr>
<td>Constant</td>
<td>368.521</td>
<td>370.627</td>
<td>399.749</td>
<td>275.719</td>
<td>270.816</td>
<td>383.100</td>
<td>378.920</td>
<td>383.100</td>
</tr>
<tr>
<td>R²</td>
<td>0.564</td>
<td>0.562</td>
<td>0.282</td>
<td>0.283</td>
<td>0.475</td>
<td>0.477</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Notes</td>
<td>Full sample</td>
<td>Full sample</td>
<td>Swiss sample</td>
<td>Swiss sample</td>
<td>Full sample</td>
<td>Full sample</td>
<td>Swiss sample</td>
<td>Swiss sample</td>
</tr>
</tbody>
</table>

Further Notes: The four leftmost columns show results in which Switzerland’s history is considered one of non-compulsory voting; the four rightmost columns take Switzerland to have practised compulsory voting between the years 1904 and 1974. The absolute values of heteroscedasticity-robust t-statistics are given within parentheses.

Johnson on www.SwissInfo.ch (http://www.swissinfo.ch/directdemocracy/democratic—the-canton-where-voting-is-compulsory/38299724), in which Schaffhausen’s cantonal chancellor mentions that “[i]t is relatively easy to be excused from voting. Accepted reasons, which are not checked, include holiday or illness, can be noted on the returned polling card. What is more, there is still three days’ grace to hand the card in, without an excuse, after voting day.”
III – MAIN RESULTS

The regression model that is of interest in what follows retains the definitions of variables introduced in the preceding section and is of the following type:

\[
\text{turnout} = \text{constant} + \beta \times \text{number of concurrent referenda} + \gamma \times \text{controls} + \epsilon
\]

where \( \epsilon \) is an idiosyncratic variable. Table II summarizes ordinary-least-squares results with voter turnout (in percentage points) as the dependant variable, with and without compulsory voting (the four right- and the four leftmost columns, respectively), and with standard control variables described in the preceding section. The sample is somewhat smaller than in Figure I because 25 observations from New Zealand lack data on the margin of victory. The constant is very high because of the \( \text{Year of Vote} \) control; with the first referendum in the sample held in 1889, the more relevant constant would be around 75 or 80 (e.g.: constant-0.16*1889=constant-302.24). The coefficient on \( \text{Year of Vote} \) is negative and statistically different from zero everywhere, indicating a downward trend in voter turnout over time. The coefficient on \( \text{Victory Margin} \) corroborates the established idea (e.g., Fauvelle-Aymar and François, 2006), that a narrower referendum attracts more voters (due to the greater chance of being pivotal, or perhaps due to more intense campaigning efforts).

More interestingly, the coefficient on the number of referenda is always negative. A squared term is also included, since the strong hint from the figure introducing this article is that turnout declines most strongly between the first and second referendum. Consistent with this observation, the squared term is always positive. To complement these two variables, a dummy, \( \text{Two or More Referenda} \), is also constructed, which groups every voting day of more than one referendum together and takes the value one for multiple concurrent referenda.

These results show that higher turnout is not associated with “cheaper” voting, in the ways standard theories of turnout define “cheap”. So far, what is less clear is whether there is a move in the opposite direction that is statistically different from zero. The specifications in Table II which assume Switzerland’s history of compulsory voting to be relevant uniformly show economically and statistically much less significant results than do the specifications which ignore compulsory Swiss voting, with the full sample being more strongly affected.

As hinted in the preceding section, there are reasons to believe that Swiss compulsory voting was never a significant factor in influencing Swiss turnout rates. A comparison with Australia and Italy, the two other countries in the sample in which compulsory voting

Table III: The Impact of Compulsory Voting on Turnout across Countries

<table>
<thead>
<tr>
<th></th>
<th>Non-Swiss</th>
<th>Italy</th>
<th>Australia</th>
<th>Switzerland</th>
</tr>
</thead>
<tbody>
<tr>
<td>( N )</td>
<td>43</td>
<td>21</td>
<td>22</td>
<td>279</td>
</tr>
<tr>
<td>Year of Vote</td>
<td>-0.042</td>
<td>-0.392</td>
<td>0.074</td>
<td>-0.164</td>
</tr>
<tr>
<td>(0.668)</td>
<td>(1.786)</td>
<td>(1.398)</td>
<td>(7.059)</td>
<td></td>
</tr>
<tr>
<td>Victory Margin (in percentage points)</td>
<td>-0.374</td>
<td>-0.118</td>
<td>-0.659</td>
<td>-0.208</td>
</tr>
<tr>
<td>(1.977)</td>
<td>(0.386)</td>
<td>(3.212)</td>
<td>(2.231)</td>
<td></td>
</tr>
<tr>
<td>Australia</td>
<td>18.000</td>
<td>–</td>
<td>–</td>
<td>–</td>
</tr>
<tr>
<td>(3.468)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(8.386)</td>
<td>(3.437)</td>
<td>(5.132)</td>
<td>(0.960)</td>
<td></td>
</tr>
<tr>
<td>Two or More Referenda</td>
<td>-6.592</td>
<td>-6.915</td>
<td>-7.975</td>
<td>-2.410</td>
</tr>
<tr>
<td>(2.185)</td>
<td>(1.428)</td>
<td>(2.015)</td>
<td>(1.429)</td>
<td></td>
</tr>
<tr>
<td>Constant</td>
<td>134.178</td>
<td>829.888</td>
<td>64.848</td>
<td>373.282</td>
</tr>
<tr>
<td>(1.077)</td>
<td>(1.903)</td>
<td>(0.655)</td>
<td>(8.152)</td>
<td></td>
</tr>
<tr>
<td>R²</td>
<td>0.852</td>
<td>0.751</td>
<td>0.907</td>
<td>0.286</td>
</tr>
</tbody>
</table>

Note: The dependant variable is voter turnout. The first column groups Australia and Italy together. Absolute values of heteroscedasticity-robust t-statistics are within parentheses.
voting is or has been practised, reveals compulsory voting to have essentially no impact in Switzerland, as seen in Table III. While the compulsory voting-dummy is associated with increases in voter turnout around twenty-five percentage points in Italy and Australia and more for them together, the corresponding figure for Switzerland is below two (some of the common findings for election turnout - e.g. Blais, 2000 - suggest about ten percentage points), and is statistically indistinguishable from zero.

Table IV: Non-Swiss within-country OLS Results

<table>
<thead>
<tr>
<th></th>
<th>N=22</th>
<th>N=22</th>
<th>N=25</th>
<th>N=25</th>
<th>N=21</th>
<th>N=21</th>
<th>N=36</th>
<th>N=104</th>
<th>N=104</th>
</tr>
</thead>
<tbody>
<tr>
<td>Year of Vote</td>
<td>0.067</td>
<td>0.072</td>
<td>-0.276</td>
<td>-0.277</td>
<td>-0.402</td>
<td>-0.392</td>
<td>-0.017</td>
<td>-0.059</td>
<td>-0.061</td>
</tr>
<tr>
<td></td>
<td>(1.278)</td>
<td>(1.411)</td>
<td>(3.117)</td>
<td>(3.155)</td>
<td>(1.672)</td>
<td>(1.786)</td>
<td>(0.269)</td>
<td>(1.178)</td>
<td>(1.189)</td>
</tr>
<tr>
<td>Victory</td>
<td>-0.595</td>
<td>-0.655</td>
<td>-0.289</td>
<td>-0.284</td>
<td>0.037</td>
<td>-0.118</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Margin (%)</td>
<td>(4.499)</td>
<td>(3.136)</td>
<td>(1.323)</td>
<td>(1.185)</td>
<td>(0.129)</td>
<td>(0.386)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Voting</td>
<td>(5.260)</td>
<td>(5.085)</td>
<td></td>
<td></td>
<td>(3.553)</td>
<td>(3.437)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Concurrent Parl.</td>
<td>-1.233</td>
<td>-0.577</td>
<td></td>
<td></td>
<td>21.621</td>
<td>11.997</td>
<td>11.468</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Election</td>
<td>(0.548)</td>
<td>(0.268)</td>
<td></td>
<td></td>
<td>(3.340)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>New Zealand</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>10.943</td>
<td>10.522</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ireland</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>-5.203</td>
<td>-5.860</td>
<td>(0.828)</td>
<td>(0.897)</td>
<td></td>
</tr>
<tr>
<td>Italy</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>-17.961</td>
<td>-17.402</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Number of Referenda</td>
<td>-11.738</td>
<td>-4.181</td>
<td>-5.099</td>
<td>-0.202</td>
<td>-4.745</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Squared Referenda</td>
<td>(2.003)</td>
<td></td>
<td></td>
<td></td>
<td>(0.158)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>1.742</td>
<td>0.656</td>
<td>0.514</td>
<td>0.501</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Two or More Referenda</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Constant</td>
<td>-43.166</td>
<td>-60.587</td>
<td>610.794</td>
<td>609.903</td>
<td>848.825</td>
<td>829.888</td>
<td>97.856</td>
<td>180.861</td>
<td>181.084</td>
</tr>
<tr>
<td></td>
<td>(0.461)</td>
<td>(0.638)</td>
<td>(3.445)</td>
<td>(3.512)</td>
<td>(1.781)</td>
<td>(1.903)</td>
<td>(0.771)</td>
<td>(1.837)</td>
<td>(1.807)</td>
</tr>
<tr>
<td>R²</td>
<td>0.908</td>
<td>0.908</td>
<td>0.234</td>
<td>0.234</td>
<td>0.779</td>
<td>0.751</td>
<td>0.452</td>
<td>0.752</td>
<td>0.742</td>
</tr>
</tbody>
</table>

Notes: The dependant variable is voter turnout. Heteroscedasticity-robust t-statistics (in absolute values) are in parentheses. It should be borne in mind that all of New Zealand’s three instances of multiple concurrent referenda were doublets, making the squared referenda term equal to the referenda term and one of them redundant. Wherever concurrent parliamentary elections are at least ten per cent of the sample they have been included as regressors, though they do not qualitatively alter the results.

Estimates excluding the Swiss sample are exhibited in Table IV. Since a large share of the sample comes from Switzerland, many of Table IV’s columns should be taken with a grain of salt as the samples get rather small. Nevertheless, all the subsamples save that for New Zealand (where only three concurrent referenda have ever been held, all doublets) again indicate declining turnout as voting becomes “cheaper”\textsuperscript{11}. An unreported regression model with the full sample and Victory Margin as a regressor has fewer observations from New Zealand but yields similar results.

\textsuperscript{11} Since small samples can induce bias in robust standard errors, additional but unreported regressions were run with non-robust t-statistics, whose magnitudes are nevertheless very close to the ones reported in Table IV, increasing somewhat for Australia and Ireland, and decreasing for Italy and the full non-Swiss sample.
The results so far suggest that voters have a tendency to eschew “quantity discounts” in referenda, which is puzzling considering their popularity elsewhere. Either there is something about the referenda that share a voting day, or there is something about the sharing itself, that lowers turnout. If the first case is true, it would imply that the present findings are due to weak identification. In the second case, voters, in contrast to established theories of turnout, are put off by the presence of multiple referenda. A post-referendum survey that has been conducted in Switzerland since 1981 asks roughly 1,000 individuals, both voters and abstainers, about the importance for them personally of the most recent instance of direct democracy. These data suggest that it is indeed the sharing itself that reduces turnout.

Importance is here measured by the fraction of respondents rating a referendum as maximally important on a ten-point scale (using the average score instead does not greatly alter the results). In cases of multiple concurrent referenda, the most important one of them serves as the rating for all, since they are held on the same day and one can just ignore those referenda one deems less important (the correlation coefficient between this measure and the number of concurrent referenda is 0.211). Since individuals may find different referenda important, this measure understates the aggregated importance for the batch of concurrently held referenda relative to stand-alone referenda.

Table V shows OLS estimates for the subsample for which these post-referendum surveys were available (compulsory voting is no longer an issue here, since these data start in the 1980’s). As can be seen, for this particular subsample, more ballots do not appear to lower turnout without controlling for the perceived personal importance of the referendum, but introducing this control restores the negative coefficients, even with the aforementioned bias for multiple referenda. At only 97 voting days, the sample size is small and the coefficients have p-values somewhat larger than 0.2 for the variables Number of Referenda and Referenda Squared, and around 0.4 for the dummy.

Note: All observations are from Switzerland between the years 1981 and 2014, which are all the referenda for which data on perceived importance are available. The absolute values of the heteroscedasticity-robust t-statistics are given within parentheses.

---

The survey is part of a series called “VOXIT” and can be found online: http://forsnesstar.unil.ch/webview/index.jsp.
However, controlling for importance does not weaken the conclusion that concurrent ballots do not yield higher turnout and likely reduce it.

A different way of holding constant the quality of referenda is to control for the issues they deal with. The database stores this information in a rough but usable fashion by attaching key words to each referendum. In Table VI, the five most frequently occurring issues (there is a total of 161 distinct issues), listed in the table, are used as control variables. Somewhat surprisingly, none of the five most frequently occurring issues is voted on concurrently with another of the five in Switzerland, indicating that the authorities wish to keep separate the most commonly voted-on issues; together they constitute 121 voting days out of 279 (the number of different issues is otherwise quite closely correlated with the number of concurrent referenda with a correlation coefficient of around 0.8). Each of these variables has a mean of around three and a median of two with respect to the total number of referenda held concurrently on the day.

Again, these controls fail to find evidence that voter turnout should rise with the number of concurrent referenda. In unreported alternative specifications with as many as the 15 most frequently occurring issues, the coefficients do not change though the number of variables somewhat reduces the statistical significance. Notice, however, that while the issues do capture a similarity of topics of referenda, it remains possible that referenda of the same topic are nevertheless perceived as unequally important.

The main reason for supposing that uninteresting referenda should be grouped together is that this would save the authorities having to organize a referendum for which there will be very little turnout. The results fail to corroborate such an “economizing” hypothesis. Another fact also argues against it; between 1889 and 2014, the average

<table>
<thead>
<tr>
<th>Year of Vote</th>
<th>-0.173</th>
<th>-0.170</th>
<th>-0.174</th>
<th>-0.176</th>
<th>-0.174</th>
<th>-0.171</th>
<th>-0.176</th>
<th>-0.177</th>
</tr>
</thead>
<tbody>
<tr>
<td>(7.329)</td>
<td>(7.325)</td>
<td>(7.534)</td>
<td>(7.529)</td>
<td>(7.415)</td>
<td>(7.461)</td>
<td>(7.649)</td>
<td>(7.607)</td>
<td></td>
</tr>
<tr>
<td>Victory</td>
<td>-0.201</td>
<td>-0.203</td>
<td>-0.200</td>
<td>-0.200</td>
<td>-0.201</td>
<td>-0.203</td>
<td>-0.197</td>
<td>-0.200</td>
</tr>
<tr>
<td>(2.131)</td>
<td>(2.140)</td>
<td>(2.064)</td>
<td>(2.069)</td>
<td>(2.150)</td>
<td>(2.160)</td>
<td>(2.073)</td>
<td>(2.078)</td>
<td></td>
</tr>
<tr>
<td>Number of Referenda</td>
<td>-2.769</td>
<td>-2.650</td>
<td>-3.067</td>
<td>-3.039</td>
<td>- - -</td>
<td>- - -</td>
<td>- - -</td>
<td>- - -</td>
</tr>
<tr>
<td>(1.838)</td>
<td>(1.734)</td>
<td>(2.011)</td>
<td>(1.956)</td>
<td>- - -</td>
<td>- - -</td>
<td>- - -</td>
<td>- - -</td>
<td></td>
</tr>
<tr>
<td>Referenda Squared</td>
<td>0.299</td>
<td>0.310</td>
<td>0.359</td>
<td>0.348</td>
<td>- - -</td>
<td>- - -</td>
<td>- - -</td>
<td>- - -</td>
</tr>
<tr>
<td>(1.692)</td>
<td>(1.688)</td>
<td>(1.991)</td>
<td>(1.928)</td>
<td>- - -</td>
<td>- - -</td>
<td>- - -</td>
<td>- - -</td>
<td></td>
</tr>
<tr>
<td>Two or More Referenda</td>
<td>-3.024</td>
<td>-2.596</td>
<td>-2.833</td>
<td>-2.954</td>
<td>- - -</td>
<td>- - -</td>
<td>- - -</td>
<td>- - -</td>
</tr>
<tr>
<td>(1.763)</td>
<td>(1.516)</td>
<td>(1.666)</td>
<td>(1.724)</td>
<td>- - -</td>
<td>- - -</td>
<td>- - -</td>
<td>- - -</td>
<td></td>
</tr>
<tr>
<td>Indirect</td>
<td>2.765</td>
<td>3.309</td>
<td>3.621</td>
<td>3.768</td>
<td>2.703</td>
<td>3.218</td>
<td>3.484</td>
<td>3.662</td>
</tr>
<tr>
<td>(1.329)</td>
<td>(1.734)</td>
<td>(1.713)</td>
<td>(1.764)</td>
<td>(1.297)</td>
<td>(1.538)</td>
<td>(1.646)</td>
<td>(1.702)</td>
<td></td>
</tr>
<tr>
<td>Taxation (27)</td>
<td>-3.889</td>
<td>-3.454</td>
<td>-3.621</td>
<td>-3.574</td>
<td>-3.161</td>
<td>-3.343</td>
<td>- - -</td>
<td>- - -</td>
</tr>
<tr>
<td>Environmental Policy (26)</td>
<td>- - -</td>
<td>4.099</td>
<td>4.179</td>
<td>- - -</td>
<td>- - -</td>
<td>- - -</td>
<td>- - -</td>
<td>- - -</td>
</tr>
<tr>
<td>Policy (26)</td>
<td>(2.010)</td>
<td>(1.781)</td>
<td>(1.825)</td>
<td>(1.854)</td>
<td>(1.648)</td>
<td>(1.708)</td>
<td>- - -</td>
<td>- - -</td>
</tr>
<tr>
<td>Immigration</td>
<td>- - -</td>
<td>4.099</td>
<td>4.179</td>
<td>- - -</td>
<td>- - -</td>
<td>- - -</td>
<td>- - -</td>
<td>- - -</td>
</tr>
<tr>
<td>Fundamental Rights (22)</td>
<td>- - -</td>
<td>2.047</td>
<td>- - -</td>
<td>- - -</td>
<td>- - -</td>
<td>- - -</td>
<td>- - -</td>
<td>- - -</td>
</tr>
<tr>
<td>Status of Women (22)</td>
<td>- - -</td>
<td>0.410</td>
<td>- - -</td>
<td>- - -</td>
<td>- - -</td>
<td>- - -</td>
<td>- - -</td>
<td>- - -</td>
</tr>
<tr>
<td>Constant</td>
<td>395.115</td>
<td>388.864</td>
<td>397.845</td>
<td>400.122</td>
<td>394.311</td>
<td>388.567</td>
<td>398.203</td>
<td>399.994</td>
</tr>
<tr>
<td>(8.471)</td>
<td>(8.493)</td>
<td>(8.720)</td>
<td>(8.699)</td>
<td>(8.513)</td>
<td>(8.587)</td>
<td>(8.778)</td>
<td>(8.726)</td>
<td>- - -</td>
</tr>
<tr>
<td>R2</td>
<td>0.285</td>
<td>0.293</td>
<td>0.301</td>
<td>0.303</td>
<td>0.286</td>
<td>0.292</td>
<td>0.299</td>
<td>0.302</td>
</tr>
</tbody>
</table>

Note: All observations from Switzerland. The issues controlled for are official descriptions of what a referendum is about, and are the issues that occur most frequently (their respective number of occurrences are given within parentheses). The absolute values of the heteroscedasticity-robust t-statistics are given within parentheses.
number of days on which referenda were held in Switzerland was 2.258 per year. If the 
economizing hypothesis were true, it seems it should be easy to economize much more 
by reducing the mean well below two. After all, if many referenda are not too important, 
they could wait half a year longer to be voted on and save significantly on the costs of 
arranging referenda.
IV – DISCUSSION

The evidence presented in the preceding section indicates that voter turnout does not increase in the number of concurrent referenda, and gives some reason to suppose that the actual relationship is the opposite. The “obvious” explanation that unimportant referenda may be grouped together does not fare well in the subsamples for which citizens’ perceived importance, or the topics of the referenda, can be added as a control variable. This section examines some potential reasons that might explain why established theories fare so badly in these data.

One explanation says that media space is scarce and that much information pertinent to the ballots consequently does not reach voters when many concurrent referenda are approaching; news reports on different referenda “crowd out” one another. This explanation is consistent with the fact that the rate of decline in turnout appears to peter out the more numerous are the referenda, akin to how one referendum might (depending on how space is allocated) lose half of its media space ahead of a double-referendum, whereas twin referenda both lose a third of their space if the twin turns into a triplet, etc. However, the media-based explanation is at odds with the finding that turnout generally increases for multiple concurrent elections, where similar media constraints ought to apply. In addition, it is not at all clear why media space would not expand when, controlling for their perceived importance, many referenda are approaching.

Some might argue the opposite: that voters face “information overload” in the presence of multiple concurrent referenda, but then one must also argue that voters somehow do not realize that they can “roll-off” (i.e. vote on some of the concurrent referenda, but abstain on those ballots they deem less interesting). One must also explain why information overload should happen already ahead of twin referenda and then proceed at a slower rate for triplets, etc., when the rate of decline in turnout should more likely increase in the number of concurrent referenda, if voters really faced too much information.

Among more established theories of voter turnout, none is supported by the present findings. Indeed, even if turnout had been found to be clearly unaffected by the number of concurrent referenda, established theories would still have been at a loss for an explanation. Expressive theories of voting predict that the greater variety in issues associated with more numerous referenda being held concurrently should attract a similarly greater variety of people interested in expressing approval or disquiet - or simply “good citizenship” or otherwise an attachment to some outcome, which does not seem to happen in the data.

Theories of instrumental voting also predict increasing turnout in the number of concurrent referenda. This is because the benefit derived if an individual’s preferred option wins likely varies across referenda, so if more referenda are held concurrently the chances increase that one of them should be very important. Additionally, the ability to vote for several proposals for the same shoe-leather cost should also raise turnout.

Information-based theories of voting also fare poorly. Feddersen and Pesendorfer (1996; 1999), and McMurray (2013) argue that uninformed voters abstain in order to strategically allow better-informed citizens to make choices for them. But on days of multiple referenda, more people ought to be informed about at least one issue that is up for vote than on days of just one referendum. If the strategic incentive were a significant factor in getting out the vote, one would expect different concurrent referenda to attract somewhat different groups. This should increase turnout though many blank votes might be cast.

13 Relatedly, Stadelmann and Torgler (2013) find that a parliamentary recommendation on how to vote has greater influence on the outcome of a referendum when there are more of them held concurrently, which does indicate a desire to economize on information over the whole set of concurrent referenda.
A different strand of voting theory deviates from methodological individualism by holding that people vote out of a sense of “ethical” (rule-utilitarian) obligation to their group (e.g. Harsanyi, 1980; Coate and Conlin, 2004; Feddersen and Sandroni, 2006). Essentially, theories of this kind argue that voting decisions are based on individuals’ reaching an optimal voting probability when the objective is to maximize group welfare. However, an individual, call him A, should revise upwards this probability in the presence of additional referenda being voted on concurrently, since the influx of voters due to the additional referenda means that a greater share of voters do not care about A’s issue (they are there for other ballots). Not knowing which way these other voters will vote on the ballot that he cares about, A sees the influx as an addition of noise to the outcome. This makes the outcome less certain and therefore makes his vote more likely to be pivotal.

If established theories of voting cannot explain the present findings, what can? Every economic explanation of turnout must in some way be traceable to the “calculus of voting” (Downs, 1957; Becker, 1958; Tullock, 1967), which essentially takes the individual as the unit of analysis and asks what incentives he faces. There are some features of the data presented in this article which cautiously suggest that theories of turnout be amended by a version of a “social incentives”-type explanation (e.g. Gerber et al., 2008; Funk, 2010; DellaVigna et al., 2014). The following provides the outline of such a model.

1. An issue that is up for vote will interest certain groups in the sense that they have high stakes in the outcome.
2. Group formation is exogenous to referenda.
3. Within groups, peer discipline pressures members to vote, although individual members would prefer not to vote since voting is costly and unlikely to affect outcomes. If a member is found not to have voted, the discoverer must punish him or be punished himself if found out. Punishment is apt to be informal and idiosyncratic, e.g. cancelled squash games, a moratorium on dinner invitations, or other ostracism.
4. The probability that more groups will have a stake in the outcome of a referendum increases in the number of referenda. Having multiple group memberships enable other groups to carry out the costly punishment, giving rise to a “prisoner’s dilemma” when all punishers want to free ride on other groups’ punishment. If punishers do not know whether a suspected non-voter also belongs to other groups, the probability that he does increases in the number of groups involved in a voting day.

Steps 1 and 2 above are innocuous enough; ceteris paribus, it is plain that more groups are mobilized when more issues are up for vote. Nor is the assumption of exogenous group formation heroic; even in Switzerland, where referenda are legion, there are hardly ever more than four days on which referenda are held per year. Not every group will be interested in every referendum, and the ability to predict what future issues will be voted on when forming new social ties is profoundly unrealistic. This also rules out individuals’ exerting effort to join multiple groups purely in the hope of avoiding punishment.

---

14 There may be special circumstances in some referenda which allow A to deduce the impact of the influx of voters, but there is no reason to suppose this to hold true very frequently, and for it to have an impact on turnout it would have to be that the added voters, who are in the same position as is A vis-à-vis his ballot, make an inference which causes them, too, to abstain from voting. To explain the data, individuals would have to consistently reason in this way, which seems highly unlikely.
As for Step 3, peer discipline within groups has been found to work to establish pro-social behaviour on the basis that non-punishment is cause for punishment when peers keep auditing one another indefinitely (see Ostrom, 1990, and Levine and Modica, 2016). Step 4 is sure to hold when multiple group membership is hard to ascertain, but even when this does not apply one can nevertheless argue that other groups should carry out the costly punishment. In either case, a non-voter has an easier time avoiding peer pressure the more groups have a stake in the outcomes of a voting day.

For those who belong to at least one group, voting thus yields utility of the form

\[ U_i = \begin{cases} 
  c_i + b_i, & \text{if voting} \\
  p/f(g_i), & \text{if abstaining} 
\end{cases} \]

where \( b_i > 0 \) is the individual’s benefit of voting (instrumental, expressive, or other), \( c_i < 0 \) is the cost of voting (time, shoe-leather, general dreariness of politics, etc.), \( p < 0 \) is the expected punishment for abstention, and \( g_i \) is the number of eligible punisher-groups (i.e. groups in which \( i \) is a member), where \( f(g) = g^{-100p} \) (recall that \( p < 0 \), so the derivative would indeed be strongly positive), since belonging to only two groups suffices to enable the “prisoner’s-dilemma” outcome. Thus, an individual votes if \( c_i + b_i > p/f(g_i) \).

For discovering punishment, the analogous decision is

\[ U_j = \begin{cases} 
  h_j(g_i), & \text{if punishing} \\
  p/h_j(g_i), & \text{if abstaining} 
\end{cases} \]

for punisher \( j \), where \( h_j(g_i) \) is a function that increases in the number of potential groups of which the suspected non-voter \( i \) is a member, whose first and second derivatives are both positive. Otherwise the notation is the same as above.

Notice that individuals may vote even if \( c_i + b_i \ll 0 \), if the expected punishment is sufficiently strong. Those who enjoy voting will do it eagerly irrespective of how high the punishment is. Those who dislike voting will do it grudgingly when the punishment is severe enough. Since the benefits \( b_i \) and costs \( c_i \) are defined generally, the present voting calculus can be seen as extending upon established theories of turnout that compare costs and benefits associated with voting. Essentially, the only thing that is done here is to add \( p/f(g_i) \) as a cost of not voting to the individual’s “calculus of voting”. It could be appended to an established turnout model, such as the classic Downsian calculus, whose predictions would then remain unchanged on the proviso that the number of concurrent referenda be held constant.

A numerical example may prove illustrative: if a group of five has an interest in the outcome of a referendum, and there is another group of five, three of whom also belong to the first group, that is concerned with a second referendum, the double referendum would mobilize two groups, or seven individuals. If these individuals know that the group-level free-riding described above will let them off the hook, only four of the seven individuals will vote (two from each group), and the intersection (the three dual-membership individuals) will abstain. Two stand-alone referenda, by contrast, would have turned out five people out of seven. More generally, the first set of voters may be thought of as those who would vote on a day of \( n - 1 \) concurrent referenda, with the second set being the group mobilized by the \( n^{th} \) referendum. As long as the intersection is (believed to be) greater than half of the influx of voters, adding a concurrent referendum will decrease turnout. This implies that the present explanation is consistent with various possible rates of decline in turnout as the number of concurrent referenda increases, and is illustrated in Figure II.
This explanation is consistent with the observation that concurrent *elections* increase turnout *if* the added election does not bring additional groups into play (or does, but without sufficient overlap with other groups). Or, to put the condition in other words, if groups (farmers, motorists, teachers, homosexuals, etc.) are about as concerned with national policy as with state or local policy, or if a main motivation to vote is partisanship or ideological affiliation. In this case, there would be no added reason to abstain from voting, and the traditional reasons for turnout to increase would apply.

If the social-incentives-and-group-based-free-riding explanation is correct, turnout should also decline by less in places and at times when turnout for referenda is lower in general. This is because the smaller is one group (or set of groups) that “care” about an issue up for vote, the lower are the chances of overlap with added groups; if there is but one group with a stake in the outcome per referendum, chances are higher that they will overlap in concurrent referenda the bigger they are. Some suggestive but weak evidence that this might happen is presented in Table VII.

Assuming that turnout varies over time for reasons unrelated to the proposed explanation, periods of unusually low turnout should have less negative, or even positive, coefficients on the multiple-referenda variables than other times. Similarly, the coefficients should be most strongly negative in times of unusually high turnout. Table VII shows average turnout for stand-alone referenda per time period so that turnout is comparable over time. In Switzerland, turnout has been at its lowest point in the present century (39.82 per cent), which is the only period when the number of concurrent referenda is positively associated with turnout. The three periods with highest turnout also have the strongest negative effects on turnout. On the other hand, the period between 1961 and 2000 averaged ten percentage points lower turnout than that between 1941 and 1960, yet has much stronger negative effects on turnout from adding concurrent referenda to a voting day.

Notice that the sign on the squared number of referenda is reversed in the 21st century. If the chances of overlap between groups are lower the smaller are the groups mobilized to vote, then one may think of the largely disjoint groups mobilized to vote on a twin referendum as constituting one group with which a different group may intersect if the twin turned into a triplet. This means that chances of significant overlap increase in the number of concurrent referenda, which should make the squared term negative in times of unusually low turnout. This is consistent with what has happened since the year 2000. However, in light of the small samples which result from the temporal partition, and given the strong assumption of exogenous variations in turnout, these are only weak indications.

A different check on the social-incentives-and-group-based-free-riding hypothesis is to compare the magnitude of the negative effect on turnout of multiple concurrent referenda across countries. As was shown above, turnout is far higher in Australia and New Zealand than in any of the other countries, which average rather similar turnout

**Figure II:** The effect of an additional referendum increases turnout (grey-shaded areas) if the (perceived) overlap between mobilized groups is slight (left), but not if great (right).
rates. The regression results shown in Table IV revealed the decline in turnout from multiple concurrent referenda to be greatest in Australia, in line with the hypothesis, but then no clear relationship can be found with the other countries and New Zealand disagrees sharply. The small number of observations for these countries makes it difficult to say more.

Table VII: Time-Segmented Swiss Referenda Regressions

<table>
<thead>
<tr>
<th>Interval</th>
<th>1889-00</th>
<th>1901-20</th>
<th>1921-40</th>
<th>1941-60</th>
<th>1961-80</th>
<th>1981-00</th>
<th>2001-14</th>
</tr>
</thead>
<tbody>
<tr>
<td>Observations</td>
<td>N=20</td>
<td>N=21</td>
<td>N=36</td>
<td>N=48</td>
<td>N=53</td>
<td>N=57</td>
<td>N=44</td>
</tr>
<tr>
<td>Avg. Turnout</td>
<td>60.74</td>
<td>57.38</td>
<td>66.70</td>
<td>53.24</td>
<td>43.48</td>
<td>44.72</td>
<td>39.82</td>
</tr>
<tr>
<td>Victory</td>
<td>0.480</td>
<td>-0.469</td>
<td>-0.498</td>
<td>-0.283</td>
<td>-0.140</td>
<td>-0.039</td>
<td>-0.038</td>
</tr>
<tr>
<td>Margin</td>
<td>(2.655)</td>
<td>(1.364)</td>
<td>(1.299)</td>
<td>(0.787)</td>
<td>(1.631)</td>
<td>(0.227)</td>
<td>(0.216)</td>
</tr>
<tr>
<td>Compulsory</td>
<td>-5.602</td>
<td>-1.216</td>
<td>-0.464</td>
<td>-0.095</td>
<td>-0.464</td>
<td>-0.038</td>
<td>-0.038</td>
</tr>
<tr>
<td>Referenda</td>
<td>(1.098)</td>
<td>(0.698)</td>
<td>(0.158)</td>
<td>(1.961)</td>
<td>(0.022)</td>
<td>(1.196)</td>
<td>(0.945)</td>
</tr>
<tr>
<td>Squared</td>
<td>7.704</td>
<td>1.953</td>
<td>0.082</td>
<td>3.067</td>
<td>_</td>
<td>0.743</td>
<td>0.762</td>
</tr>
<tr>
<td>Constant</td>
<td>74.938</td>
<td>68.607</td>
<td>69.562</td>
<td>85.205</td>
<td>54.693</td>
<td>47.925</td>
<td>48.526</td>
</tr>
<tr>
<td>R²</td>
<td>0.226</td>
<td>0.165</td>
<td>0.186</td>
<td>0.188</td>
<td>0.018</td>
<td>0.022</td>
<td>0.022</td>
</tr>
</tbody>
</table>

Alternative results using the binary variable Two or More Referenda to capture the effect of multiple referenda

| Victory       | 0.428   | -0.472  | -0.487  | -0.281  | -0.140  | -0.057  | -0.055  | -0.199  | -0.327   |
| Margin        | (2.109) | (1.461) | (1.407) | (0.799) | (1.631) | (0.363) | (0.323) | (0.757) | (2.913)  |
| Compulsory    | -5.096  | -1.473  | -0.617  | -0.066  | -1.492  | -1.597  | -4.968  | -5.308  |
| Referenda     | (0.750) | (0.713) | (0.856) | (2.000) | (0.022) | (0.502) | (0.294) | (0.719) | (2.026)  |
| Constant      | 53.936  | 61.534  | 66.250  | 70.205  | 54.693  | 44.168  | 44.345  | 47.871  | 44.439   |
| R²            | 0.196   | 0.165   | 0.184   | 0.190   | 0.018   | 0.008   | 0.008   | 0.041   | 0.287    |

Note: The dependant variable is again voter turnout. Average turnout is mean voter turnout for stand-alone referenda in the time period in question. Note that the squared referendum variable is missing for the period 1941-1960, as all the multiple referenda held in Switzerland during that time were twins. Absolute values of heteroscedasticity-robust t-statistics are given within parentheses.
V – CONCLUSION

The main contribution of this article is to document some highly puzzling evidence on voter turnout. A second contribution is to relate the evidence to some plausible and established theories of voter turnout, and see how they fare. Since they fare poorly, a third and somewhat more tentative contribution is to suggest a variation on social incentives to vote as a potential explanation which (so far) seems broadly consistent with the data as well as with somewhat more crude tests of its implications.

The estimates provided by the regression analyses in this article suggest that turnout decreases by about two or three percentage points in a twin referendum compared to a single referendum, and that the rate of decrease declines in the number of additional referenda. The same sign appears across different times and countries, although the sizes of the effects vary and are not unfailingly statistically different from zero, which, however, would still be at odds with established theories of voting.

Since established theories of voting predict that turnout increases in the number of concurrent referenda, these results are profoundly puzzling for them. A potential explanation which is broadly consistent with the present observations holds that many individuals privately find voting too costly but vote anyway because they fear even costlier intra-group discipline. If multiple referenda attract more groups, it is likely that their memberships intersect, which induces punishers to free-ride on other groups and can also make punishment costlier, thereby letting abstainers off the hook.
REFERENCES


Blais, André, To Vote or Not to Vote: The Merits and Limits of Rational Choice Theory, University of Pittsburgh Press, 2000.

Centre for Research on Direct Democracy, http://www.c2d.ch/ (the source of most data)


This paper examines restaurant meal orders and customers’ observable characteristics (such as age, sex, and body type) gathered by a waitress at a Finnish restaurant and finds that an increase in the relative heaviness of a customer compared to the table average increases the likelihood that the customer orders a light meal (heavier people otherwise eat, if anything, less salutary meals). The findings are consistent with a “signalling theory” of group dining, according to which social image rather than health influences what people choose to be seen to eat. Companions’ orders also greatly impact own orders, in line with more macro-level studies on social transmissions of obesity such as the famous ‘contagious-obesity’ study by Christakis and Fowler (2007). (JEL codes: I12; Z13; C93)

15 Many thanks to the waitress/research assistant for her excellent, high-quality data collection without which this article could not be written. Thanks also to Ola Andersson, Fredrik Carlsson, Marja-Liisa Halko, Topi Miettinen and Tuomas Nurminen for helpful discussions. The usual disclaimer applies.
I - INTRODUCTION

Using data collected by a student spending a university vacation working as a waitress, this paper examines meal orders and observable characteristics (such as age, sex, body type, and number of dining companions and their characteristics) of customers at a Finnish restaurant. It finds that meals tend to be lighter when ordered by persons of an ampler body type than the average of the company with which they dine, even though body type alone has no consistent effect on measures of meal lightness (if anything portlier guests eat less lightly). Additionally, and independently of relative body type, companions’ averages are found to greatly impact an individual’s orders.

People’s dietary decisions are an important public-health issue and how such decisions are influenced by their social environment is in need of micro-level evidence. It is generally believed that people alter their behaviour when observed compared to when not (see, e.g. Veblen, 1899; Goffman, 1958, for classical treatments; Hopkins and Kornienko, 2004; Samuelson, 2004, for recent theoretical ones), and that the sort of alteration made depends on one’s audience (Austen-Smith and Fryer, 2005; Ellingsen and Johannesson, 2008). Laboratory evidence on the importance of external motivations is found in Malmendier et al. (2014) and in Butera and Horn (2017), while Babcock et al., 1996 find that teachers strike more frequently the greater is the salary difference in comparison districts chosen by their unions and chosen by school boards. Indeed, ‘the usual basis of self-respect is the respect accorded by one’s neighbors’ (Veblen, 1899, p. 20).

A more recent but expanding body of literature has examined decision making in groups compared to individual decision making (see Charness and Sutter, 2012, for an excellent survey). The general thrust of this literature is that group decision-making brings about better outcomes, lessening the effect of self-control problems in the face of pressure to please one’s peers. Thus, field studies of peer effects on worker productivity find substantial positive effects by mere proximity to others (Falk and Ichino, 2006), and when individual payment depends on group effort (Hamilton et al., 2003). In the lab, Kocher et al. (2006) additionally find that, when the decision of whether to work alone or in a group is left to the individual, people choosing groups report being more motivated by results and payoffs.

A final strand of related literature is that on attitudes towards more heavy-set individuals. A survey on attitudes towards the obese in the general population of Germany by Hilbert et al. (2008) shows that almost a quarter of respondents reported “stigmatizing” attitudes towards the obese, as indicated by beliefs that obesity is due to choice or not strongly predicted by heredity. But what is considered obese varies by context. Burke and Heiland (2011) document that Black and White American women differ greatly in body mass indices16 (BMI). Elsewhere, these authors also report that the difference is reflected in disparate reports of ‘ideal BMI’ (2008). Relatedly, Ali et al. (2013) find that a high BMI is perceived as less attractive among White females than among their Black counterparts. Cavley (2004) finds evidence that the labour market punishes high-BMI individuals by less if they belong to high-BMI demographic groups. The present article contributes micro-level evidence from a natural setting on how restaurant-goers choose to present themselves with respect to health consciousness. If there is pressure to be seen to ‘eat right’ and this pressure is greater for the relatively portly, a substantial share of salutary victuals would be consumed mainly for signalling purposes since many consumers would not privately desire them. By foregoing a preferred but less healthful meal in favour of one that peers may consider more apposite,

16 This index is constructed by dividing one’s weight in kilos by one’s squared height in metres. Index values roughly between 19 and 25 are considered normal. A person is generally considered (severely) obese for a BMI of 30 (35) or greater.
one incurs a cost which strengthens one’s credentials as a health-conscious individual in spite of one’s body type.

In the complementarities approach to peer effects (e.g. Becker and Murphy, 2000), what is consumed in company at a restaurant is in effect the presentation of self as an individual leading an appropriate lifestyle and what is appropriate is determined by some average of the salutary status of the group as judged by their waistbands. Similarly, being thinner than one’s dinner companions provides opportunities to indulge more because one’s thinness indicates to the group that one is quite successful at leading a healthy lifestyle.

Another contribution is this paper’s finding support for peer effects more generally in the spread of dietary habits, as individual orders are found to be strongly predicted by average group orders. Most such evidence comes from more macro-level approaches. The famous ‘contagious-obesity’ study by Christakis and Fowler (2007) examines networks in Framingham, Massachusetts, by combining panel data on physical exams and lists of named friends. In a similar vein, Trogdon et al. (2008) find a positive impact of peer weight on own weight using nominated friends within schools and various techniques to account for the endogeneity of peer groups. By contrast, no study appears to have directly examined health-affecting situations and peer-influence on decisions.

The type of interaction that the present set-up can study is akin to what is known as contextual (or exogenous) interaction (Manski, 2000, p. 127; also 1993, pp. 532-533): “the propensity of an agent to behave in some way varies with exogenous characteristics of the group members”, in contrast to endogenous interactions and correlated effects, wherein the behaviour of the group and characteristics or histories common to all group members, respectively, drive behaviour. Since behaviour differs along identifiable relative characteristics within table companies over many different compositions of these dining companies, it is unlikely that common experiences or cues from other group members should be behind the observations. The finding that individual orders are strongly influenced by table average orders is more difficult to categorize as a contextual, endogenous, or correlated effect, however.

The remainder of this paper proceeds as follows. Section II describes the data and the restaurant in greater detail, as well as the methodology. Section III presents the main results on incentives to order less caloric meals and the phenomenon that individual orders are similar to average company orders. Section IV discusses issues of interpretation, exploring possible alternative interpretations for the results. This section also considers suitable theoretical settings for the empirical findings and implications for the rationality of overeating and fat taxes. Section V concludes.
II – DATA AND METHODOLOGY

While working as a waitress during a university vacation, a student collected data on restaurant-goers’ observable physical characteristics and their orders. The data were collected unbeknownst to the customers and the restaurant was a mid-level sit-down one (save for the occasional take-away pizza) in Western Finland. In this way, observable characteristics and meal orders of 477 restaurant-goers were obtained. The waitress/research assistant was not aware that this study would use her observations to examine peer effects among meal companions.

The data were collected mainly at off-peak hours around lunchtime, using special report cards designed for this purpose. Figure I displays the likeness of an English translation of one such card. The waitress/research assistant marked her estimations of age and body type with an X on the scales, and indicated race and sex as appropriate. On the reverse side, she indicated meal information for the person to whom the card refers, including potential desserts, drink orders, and whether the orderer had any special requests such as a reduced portion, different side dishes, or a vegan version of a meal. She grouped cards by table, making sure that each individual at every table from which she collected data was represented. The 477 customers sampled together represent 218 tables.

There is no caloric labelling of meals at this restaurant, but four main courses out of the eleven listed on the menu are obvious light choices. Restaurants often design their menus to include light options and this is expressly the case here. Comparisons with the 16 varieties of pizza offered are hard to make, but none of the pizzas would qualify as a light option, except for the reduced size pizzas. The main variables with respect to the menu choices which most plausibly signal a person’s body-type consciousness are those of ordering (1) a light main course or (2) not ordering chips (‘fries’) as a side dish.

These variables figure heavily in binary-choice regression models estimated in Section III. Chips have a reputation for being a less healthy choice than the restaurant’s other side-dish options: steamed potatoes, steamed vegetables, or potato wedges. One could technically request more than one side dish but this does not happen in the sample and would be highly unusual anywhere.
The sampling was carried out between 30th May and 14th July, 2016, which is the restaurant’s busy part of the year. The restaurant is in close vicinity to a theme park which is a local attraction during the data collection period. In terms of prices and clientele, the restaurant caters mainly to middle-class customers, with prices ranging from ten euro for the cheapest pizzas or a Caesar Salad, to 26 euro for a fillet of beef, with the mean price for the non-pizza menu being €18.45 (neither drink nor tip included). Of the restaurant’s eleven main courses, four are categorized (by the restaurant’s staff) as light meals. It is also possible to order a reduced-size main course, which several guests do (except when the full size is already one of the light varieties). Reduced-size meals are coded as light.

Because all these data are available for each person at every table, it is possible to compare one person’s characteristics with those of his or her companions. Thus, Relative body type is the main variable of interest, and is constructed by taking the number of millimetres on the waitress’ mark on the 48-millimetre scale that a restaurant-goer is above the mean of the corresponding marks of his or her table companions (excluding themselves). Missing data in one three-person table is why the numbers in panel (b) of the table do not quite add up. Body types reflect the number of millimetres (out of 48) from the leftmost edge on the score cards. The figures within parentheses are standard deviations in panel (a), and, in panel (b), the number accounted for by persons of relatively large body out of the unparenthesized number.

Table I lists the main summary statistics on individuals and their orders. Lone guests are somewhat heavier than are guests who dine together. Lone individuals of larger body type tend to order more chips and eat less lightly than do their thin counterparts, whereas the relatively portly eat more lightly in groups. Since lone and accompanied guests plausibly differ in mentality and other characteristics, one cannot infer from this observation that stout persons use meals as a signalling device, although it is consistent
with that interpretation. Within groups, persons of relatively large body type account for 48.56 per cent of the guests but order 57.14 per cent of the light dinners (20/35) and 37.89 per cent of the chips (36/95).

Age plausibly impacts perceived body type since fattening habits accumulate weight, causing weight to increase with age. And as health deteriorates with age, a concomitant increase in sedentary lifestyles would naturally give rise to the same relationship. The correlation coefficient between age and body type (measured in millimetres) is 0.346. Figure II displays a scatterplot with the relationship between body type (vertical axis) and age (horizontal axis) as found in the collected observations. A comparison between the sample and the general Finnish population shows no sign of bias in the sample. In a study on socioeconomic background and obesity, Rahkonen et al. (1998, p. 92) find roughly the same relationship between body type and age as the one depicted in Figure II; an increase of about ten per cent in BMI from age 20 to age 70, going from slightly above 23 to almost 27.

![Figure II: Correlation between age and body type in the sample.](image)

A different potential issue is related to a potential lack of variation in body types at tables. On the 48-millimetre body type indicator line, the mean two-person table has a maximal body-type difference of 5.62. This figure is 12.02, 9.48, 17.54, 8.88, and 21.88 for three-, four-, five-, six-, and seven-person tables, respectively.

The appendix lists the entire menu, along with which meals the restaurant considers light. Since restaurant-goers may have different ideas about what different body types should eat, these indications are necessarily somewhat fluent. However, the classifications in the appendix would go along well with those of many other restaurants.
III - RESULTS

Table II presents the results of a binary-choice regression model which uses a meal order of interest as the dependant variable, with relative body type and control variables on the right-hand side:

\[ meal \ type = constant + \beta \times \text{relative body type} + \gamma \times \text{controls} + \epsilon. \]

| Table II: Results on Groups with Different Dependant Variables for Health Consciousness |
|-----------------------------------------------|---|---|---|---|---|
| Probit estimates | Light Meal | Chips | Unhealthy |
| N=380 | N=380 | N=380 | N=380 | N=380 | N=380 |
| Man | - | -0.905 (3.765) | - | 0.237 (1.544) | - | 1.091 (4.092) |
| Absolute body type | -0.048 (2.262) | -0.040 (1.376) | -0.007 (0.370) | 2.9*10^-4 (0.012) | 0.050 (2.414) | 0.034 (1.154) |
| Age | - | 0.016 (1.881) | - | 4.5*10^-4 (1.064) | - | -0.019 (1.828) |
| Relative body type | 0.046 (2.526) | 0.046 (2.069) | -0.029 (1.754) | -0.049 (2.419) | -0.048 (2.587) | -0.049 (1.940) |
| Table’s average | - | 2.412 (1.143) | - | 1.706 (7.972) | - | 2.887 (7.788) |
| Constant | -0.183 (0.354) | -0.968 (1.368) | -0.531 (1.212) | -1.337 (2.469) | 0.218 (0.012) | -1.546 (2.105) |
| R² | 0.021 | 0.346 | 0.021 | 0.195 | 0.034 | 0.440 |
| # of 1’s in Y | 35 | 35 | 95 | 95 | 349 | 349 |

<table>
<thead>
<tr>
<th>OLS estimates</th>
<th>Light Meal</th>
<th>Chips</th>
<th>Unhealthy</th>
</tr>
</thead>
<tbody>
<tr>
<td>N=380</td>
<td>N=380</td>
<td>N=380</td>
<td>N=380</td>
</tr>
<tr>
<td>Man</td>
<td>-</td>
<td>-0.090 (3.371)</td>
<td>-</td>
</tr>
<tr>
<td>Absolute body type</td>
<td>-0.008 (1.550)</td>
<td>-0.005 (2.608)</td>
<td>-0.002 (0.260)</td>
</tr>
<tr>
<td>Age</td>
<td>-</td>
<td>0.002 (1.750)</td>
<td>-</td>
</tr>
<tr>
<td>Relative body type</td>
<td>0.007 (1.926)</td>
<td>0.005 (2.565)</td>
<td>-0.009 (1.571)</td>
</tr>
<tr>
<td>Table’s average</td>
<td>-</td>
<td>0.624 (5.299)</td>
<td>-</td>
</tr>
<tr>
<td>Constant</td>
<td>0.275 (2.047)</td>
<td>0.161 (3.281)</td>
<td>0.292 (1.742)</td>
</tr>
<tr>
<td>R²</td>
<td>0.012</td>
<td>0.321</td>
<td>0.022</td>
</tr>
</tbody>
</table>

Notes: The variable Relative body type refers to the difference in body type (measured in millimetres from the leftmost edge of the scale on the score cards) between the individual and the average of his or her table companions. Table’s average refers to what the other people at the same table ordered with respect to the meal in question; it is one if everyone ordered the meal, zero if no-one did, and in-between in proportion to the guests who did. The variable Unhealthy takes the value one if a person ordered either or both of a non-light meal or chips. Table-clustered t-statistics or z-statistics are given within parentheses. Note that the critical value of a t- or z-statistic is 2.4 at the five per cent level with a Bonferroni correction for making three tests; 2.13 at the ten per cent level, and 2.94 at the one per cent level.

Because the dependant variable is a dummy, results are presented from a probit as well as from an OLS regression for each specification. The OLS results are useful for interpreting the effects, bearing in mind that each unit change in body type corresponds
to one millimetre on the score card (and 0.4 BMI points, with the example of the 18 BMI point range). Errors are clustered at the table level since all individuals at a table are affected by errors pertaining to that table. With three main variables capturing orders’ potential signalling value, the critical values for the t- or z-statistics must be adjusted upwards. Using a Bonferroni correction, the critical value at the five per cent level rises from 1.96 to 2.4. These values are listed in the notes to each table that shows regression results.

The meal variables take the value of one if the meals were ordered and of zero otherwise, so that the sign on Relative body type is negative for fattening variables (Chips and Unhealthy), and positive for Light Meal, under the signalling hypothesis. Uniformly, the signs on the coefficients are indeed in the direction suggested by a signalling hypothesis, although the effects are not uniformly statistically indistinguishable from zero at the five per cent Bonferroni-adjusted level; in the specifications using more control variables, Chips is borderline significant at the five per cent Bonferroni-adjusted level. The same is true for Light Meal and Unhealthy in the OLS estimates, and in the probit estimates without controls. Oddly, the z-statistics also differ from the t-statistics in the regressions on Light Meal and Unhealthy, for relative as well as absolute body type.

Unreported regressions that interact Relative body type with itself (preserving the negative sign on the relatively thin individuals) additionally have interaction terms uniformly statistically indistinguishable from zero at even the ten per cent level and their coefficients are small enough that nonlinearities do not change point estimates by much until a body type is very far from the other ones at the table. Unreported regressions of the same variety as those in Table II but run only on persons thinner and bigger than their companions, respectively, also suggest that being thinner than one’s table companions is associated with increased caloric intake, the opposite of what happens for the relatively heavy individuals. These additional results both suggest a fairly linear relationship between relative body type and meal type.

Because the regression model uses relative body type along with absolute body type, an increase in a person’s body size affects both the absolute and the relative variables. Therefore, as robustness checks, additional regressions were run using largest body type and body type larger than average, in lieu of relative body type. The results are qualitatively the same as those in Table II.17

Table III displays results analogous to those of Table II, but on the lone-diner subsample only (N=93). No effect here is statistically significant at any conventional level with a Bonferroni adjustment, but the signs on Body type consistently suggest that heavier restaurant goers in this sample eat somewhat less salutary meals, the opposite of what happens for persons dining in company who are of a relatively larger body type. Again, lone eaters may differ from accompanied restaurant-goers in ways that complicate comparisons between the populations, but this tendency is at least consistent with the signalling hypothesis.

The findings on lone eaters are broadly consistent with those on people eating together. Having an absolutely larger body type in groups is associated with either eating more chips and fewer light meals, or with no change, when compared to orders of absolutely larger body types who eat alone. Thus, if an entire table company consisted of persons of a large body type, that table would order more calories than would a table of thinner individuals, but those of a relatively large body type would order fewer calories than their companions at either table.

Apart from the apparent signalling, a persistent result is that of ordering the same food as the rest of the table; the variable Table’s average in Table II is positive and statistically

17 The results of all these regressions are available upon request.
significant at even the one per cent Bonferroni-adjusted level. Whether this reflects companions’ shared initial tastes or a peer effect is unknown. At least to some extent, the data suggest that the latter effect is an active component of meal decisions. For one thing, (unreported) regressions on individual red-meat alternatives (there are five of them at the restaurant, essentially schnitzels or beef) reveal statistically significant and substantively strong coefficients in each case. On the assumption that group-formation may occur on the basis of a common preference for red meat but not on preferences for a particular kind of steak (e.g. a bovine- or porcine-based one), this is suggestive of emulation.

Furthermore, of the 125 tables with more than one person in the sample, 40 were all-(non-fish)-meat tables. Out of these 40 tables, only eleven ordered different kinds of meat. Thirty-one tables were all-pizza tables and eight of them even ordered the same exact toppings, though the restaurant offers sixteen varieties of pizza and accommodates different topping requests.

Table III: The Health Consciousness of Lone Diners

<table>
<thead>
<tr>
<th>Probit estimates</th>
<th>Light Meal</th>
<th>Chips</th>
<th>Unhealthy</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>N=93</td>
<td>N=93</td>
<td>N=93</td>
</tr>
<tr>
<td>Man</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>-1.035</td>
<td>-0.995</td>
<td>0.248</td>
</tr>
<tr>
<td></td>
<td>(2.080)</td>
<td>(1.976)</td>
<td>(0.727)</td>
</tr>
<tr>
<td>Body type</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>-0.049</td>
<td>-0.053</td>
<td>0.030</td>
</tr>
<tr>
<td></td>
<td>(1.036)</td>
<td>(1.174)</td>
<td>(1.323)</td>
</tr>
<tr>
<td>Age</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>-0.015</td>
<td>0.021</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.919)</td>
<td>(1.460)</td>
<td></td>
</tr>
<tr>
<td>Constant</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>0.519</td>
<td>0.202</td>
<td>-2.664</td>
</tr>
<tr>
<td></td>
<td>(0.464)</td>
<td>(0.174)</td>
<td>(3.622)</td>
</tr>
<tr>
<td>R²</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>0.161</td>
<td>0.155</td>
<td>0.043</td>
</tr>
<tr>
<td># of 1’s in Y</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>5</td>
<td>5</td>
<td>14</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>OLS estimates</th>
<th>Light Meal</th>
<th>Chips</th>
<th>Unhealthy</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>N=93</td>
<td>N=93</td>
<td>N=93</td>
</tr>
<tr>
<td>Man</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>-0.108</td>
<td>-0.108</td>
<td>0.050</td>
</tr>
<tr>
<td></td>
<td>(1.778)</td>
<td>(1.790)</td>
<td>(0.671)</td>
</tr>
<tr>
<td>Body type</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>-0.006</td>
<td>-0.006</td>
<td>0.006</td>
</tr>
<tr>
<td></td>
<td>(0.923)</td>
<td>(1.096)</td>
<td>(1.313)</td>
</tr>
<tr>
<td>Age</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>-0.4*10⁻⁴</td>
<td>-0.004</td>
<td>0.000</td>
</tr>
<tr>
<td></td>
<td>(0.201)</td>
<td>(1.346)</td>
<td></td>
</tr>
<tr>
<td>Constant</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>0.284</td>
<td>0.279</td>
<td>-0.178</td>
</tr>
<tr>
<td></td>
<td>(1.731)</td>
<td>(1.617)</td>
<td>(1.223)</td>
</tr>
<tr>
<td>R²</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>0.069</td>
<td>0.069</td>
<td>0.033</td>
</tr>
</tbody>
</table>

Notes: The variable Unhealthy takes the value one if a person ordered a non-light meal or chips. Heteroscedasticity-robust t-statistics or z-statistics are given within parentheses. Note that the critical value of a t- or z-statistic is 2.4 at the five per cent level with a Bonferroni correction for making three tests; 2.13 at the ten per cent level, and 2.94 at the one per cent level.
IV – DISCUSSION

One should interpret these results with a high degree of caution. For one thing, it is not certain that the waitress was skilled at estimating age and body type. There is therefore a possibility that the data are quite noisy, so that the results should be taken as provisional. Suitable tests could be given data collectors in follow-up studies in which (pictures or videos of) real people are to be judged on body type, with test performance determined by comparisons to actual ranks.

There is also no data on whether food was shared among table companions, nor information on who paid or when precisely the meals were consumed. The statistical relationships are also only borderline statistically significant with a Bonferroni adjustment. Thus, these results are merely a first and small glance at what a novel approach may yield.

To ascertain whether the observed pattern of orders is causal (from relative body type to orders), one would ideally look at the same individual, hold his appetite and emotional disposition constant, and randomly vary his companions’ characteristics. This is not feasible since people choose their dinner company. Consequently, this study should be taken to offer suggestive evidence that is consistent with a signalling theory, rather than necessarily illustrating a causal link from relative body type to food. It may, however, be difficult to think of alternative explanations which would not imply peer effects.

As an example of one alternative explanation, it could be that the relatively heavy-set individuals are influenced by their dinner companions to diet earnestly rather than signal weight consciousness. If this particular explanation were true, it would still point to peer effects but the choice of a light meal would then be socialized – wanted for its own sake rather than to serve as a signal. On the other hand, explanations based upon socialization are less simple than those based upon signalling, as the former would have to argue that those with thin dinner companions let themselves go before their preferences become socialized, rather than being socialized and of a leaner body type all the time.

The difference in meal choices between lone restaurant-goers and people in groups could reflect their situations being differently conducive to reflection and looking ahead. Rangel (2013) lists three types of dietary choices which people may make: Goal-directed (akin to rational choice), Pavlovian (automatic responses to stimuli), or Habitual (similarly unreflective decisions but independent of external stimuli). It could be that people make better or more forward-looking choices in groups and act more habitually when on their own (cf. the literature on group versus individual decisions mentioned in the introduction). This possibility is, however, inconsistent with the observation that the relatively thin eat less salutarily when in groups.

Regressions of own orders on averages of other persons at the table uniformly and strongly indicate the presence of peer effects. At least to some degree, it is likely that these are emulations of other individuals rather than a reflection of how groups form, because even within companies which order very similar main courses there is a tendency for them to be not only similar but identical. While not altogether surprising to seasoned restaurant goers, this finding points in the direction of network effects in the spread of body types in general and obesity in particular (Christakis and Fowler, 2007).

A group of people initially following identical diets will on average order victuals higher in calories if just one member increases his intake. Evidently peers can also develop healthier diets together; this merely speaks to the magnifying effect which social interactions have to make rather small initial changes (say, physically less demanding employment or cheaper calories) translate into large increases in waistlines.
If individuals have a utility function in which the utility of some consumption good \( c \) depended on certain of one's attributes, \( a \), as well as group consumption, \( s \) as captured by a utility function

\[ u = u(c, a(c), t; s), \]

where \( t \) is alternative consumption lacking a social component, then the present context would have the marginal utility of \( c \) rise with group consumption \( s \), as well as with one's attributes \( a \) up to a certain point \( a^*(c) \) whereupon consumption the marginal utility of \( c \) begins to decline. Philipson and Posner (2003) suggest a similar model in their treatment of obesity, although without social influences. If \( c \) is calories and \( a \) is one's waistband, this framework captures the pressure felt by the relatively portly members of a group of restaurant-goers, as well as the 'licence' of the relatively thin to consume more \( c \) and the network effect represented by \( s \).

It follows from this simple setting that in the presence of many peers consuming many calories, a relatively stout individual would wish to eat more from the pure effect on utility of \( c \), but would want to reduce his caloric intake due to \( c \)'s interaction with his characteristics \( a \) (his relative embonpoint). All else equal, this implies that thin individuals would wish to eat with heavier individuals, and that heavier individuals would like to avoid eating with too many thin individuals.

Assuming that the utility differences due to dining with persons of different body types are small enough relative to the other qualities a person may offer over time, the given utility function adequately represents situations such as the one described in the present paper, and indeed many situations in which the utility of consumption depends on peer consumption as well as on observable personal characteristics. In this case, whatever importance individuals place on the body types of their companions will not affect the company they choose.

Generalizing the treatment to situations other than restaurant dining, it will sometimes be possible to signal by other means than adjustments to one's consumption of \( c \). In such cases, \( a \) would be a function of both \( c \) and this other signal. For instance, cook-books with healthy recipes and gym memberships to keep on display for visitors would serve a signalling purpose if one is on the heavy side; if one's smoking or drinking is difficult to hide, various self-help products for staying off cigarettes or time spent in AA meetings can serve as a signals for smokers or alcoholics, etc.

Externally to the issue of restaurant orders, support for the signalling hypothesis suggests that ostensible concerns with caloric intake (or any behaviour regarded as a shortcoming) does not count as evidence that dieters (or consumers of self-help more generally) have 'self-control' issues, as has sometimes been argued (e.g., Gruber and Köszegi, 2001; Cutler et al., 2003, pp. 112-113). A person is rationally overweight if he sufficiently enjoys eating and dislikes exercise. Ostensible health concerns like ordering light meals among thin company enable such people to preserve more of their image. Appearances can then more easily be blamed on a funny metabolism, 'big-bonedness', or other physiological conditions. To the extent that beliefs about people's willpower influence views on 'fat taxes' and related actions, the present findings may temper enthusiasm for such taxes.

Regardless of the particular habit in question, the signal is a serviceable way to get to keep engaging in a privately cherished but socially disapproved kind of consumption. Technically, any habit that one's peers find annoying could give rise to this sort of mechanism, but addictions are likely the best candidate for such habits since non-addictive habits lack the excuse that kicking them putatively requires a degree of self-control the summoning up of which can take a long time.

The methods discussed are ways of signalling which are costly in the sense that retaining the esteem of one’s peers requires consumption decisions which are privately
disliked. Peers are more likely to perceive signals as earnest the more expensive they appear. The pressure to signal is therefore greater for persons known to be especially concerned with how they are perceived by others (akin to the distinction between intrinsic and extrinsic rewards in Bénabou and Tirole, 2006; see also Ariely et al., 2009). Whether any of these methods might apply outwith restaurants is unknown and in many cases hard to determine by simple figures, because self-help products and methods are bought by people who are commonly believed to be addicts and so if their habits do not change it could be consistent with both addition as well as with signalling.
V – SUMMARY

Observations on restaurant-goers indicate two main findings. Firstly, that having a larger body type relative to one’s table companions reduces caloric intake whereas having an absolutely large body type does not. Lone restaurant-goers have a tendency to order larger meals, and among groups light meals are ordered more often by those who are relatively heavy and less salutary meals by those who are relatively lean. The results on meal-size reductions are not uniformly statistically different from zero, however, and the relatively low number of observations makes this study a preliminary one.

Another finding in this paper is the potential network effect in the kind of meal one orders. This is stronger than the signalling effect, but to what extent it consists of emulation and to what extent it reflects network formation is unknown. Since the identification strategy in this article relies purely on control variables, both of these results are firstly to be seen as consistent with signalling and with network effects.

Generalizing from restaurant-goers, people may signal values conforming to their peers’ standards in ways other than reduced caloric intake when observed. This possibility makes ostensible attempts to end ‘addictive’ behaviours an implication of the signalling of shared values rather than a verdict on the general difficulty to kick socially disliked habits. Of course, it remains possible that failed attempts at ending addictions are explained by more than one theory.
REFERENCES


**APPENDIX: THE MENU**

### Children's menu
- Sausages: €6.00
- Minihamburger: €7.00
- Minute steak: €8.00
- Minipizza: €7.00

### Regular menu
- House's pepper steak: €25.00
- Tenderloin with chantarelle sauce: €26.00
- Beef with onions: €23.00
- Fried salmon with cream sauce: €21.00
- Schnitzel: €17.00
- House's schnitzel: €25.00
- Caesar salad: €10.00
- Beef sandwich: €14.00
- Chicken sandwich with goat cheese: €15.00
- House's hamburger platter: €14.00
- Moules-frites (mussels with chips): €13.00

(Aura cheese is a type of blue cheese, aged for six weeks in its regular variety)

### Dessert
- Chocolate cake with vanilla ice cream: €7.00
- Vanilla ice cream with a selection of sauces: €5.00
- Panna cotta: €6.00

---

**Pizza**
- Bacon: €10.00
- Tex-Mex: jalapeno, taco sauce, pepperoni: €10.00
- Vagabond: ham, shrimp: €9.50
- Mexico: kebab, onion, pepper: €11.00
- Ciccolina: chicken, pineapple: €9.50
- Columbus: kebab, béarnaise sauce: €10.50
- Manhattan: smoked salmon, shrimp, aura cheese: €11.00
- Vegetariana: pineapple, onion, pepper, olives: €9.50
- Opera: tuna, ham: €9.50
- Capricciosa: ham, mushroom: €9.50
- Pepperoni: tuna, pepperoni: €9.50
- Osolemio: tuna, shrimp: €9.50
- Americana: ham, pineapple, aura cheese: €10.00
- Quattro: ham, tuna, shrimp, mushroom: €10.50
- Nina: tuna, aura cheese, onion, pepperoni: €10.50
MARRIAGE AND BIOLOGY

Mats Ekman

Since only women can give birth, only female infidelity can put householders at risk of raising alien genetic material. This means that husbands have more to lose from their partners' infidelity and therefore have a stronger interest in deterrence than their wives do. If men earn most of the household’s market income, they are more able to withhold consumption opportunities from a spouse than if the roles were reversed, which enables them to pursue such deterrence. This paper develops a simple model to capture these incentives. It is consistent with several widely-observed phenomena and also produces novel testable implications.

(JEL codes: D10; J12; J16, J24, J31)

Iago: 'If you are so fond over her iniquity, give her patent to offend, for if it touch not you it comes near nobody.'

Othello: 'I will chop her into messes! Cuckold me?'

-William Shakespeare, Othello

\[\text{\textsuperscript{18}}\text{ Many thanks to Professors Ola Andersson, Fredrik Carlsson, Markus Jäntti, Klaus Kultti, Topi Miettinen, and Rune Stenbacka for helpful comments and suggestions.}\]
I - INTRODUCTION

This paper presents a new way of looking at households’ economic behaviour from the perspective of biology. It is a fact of biology that only women bear children and that consequently only a man can unwittingly raise offspring that is genetically alien to him. If preferences for raising offspring with shared DNA have at least some genetic component, then evolution favours men who wish to raise their own offspring, because men who do not care will not have as many (own) children. This fact has been a starting point for many treatments of unselfish behaviour by fathers towards their offspring coupled with selfishness in relation to other children19.

Assuming that there is a preference for raising one’s own offspring, it follows immediately from the aforementioned biological fact (that only women bear children) that men have more to lose from being cheated on by their wives, than have their wives in the reverse situation (although wives surely dislike being two-timed, biology implies that they do not dislike it as much as do men). Cuckolded men forego the opportunity to spend resources on raising their own offspring when they raise the children of some other man, which is detrimental to their genes. The classic law-and-economics insight that the severity of the punishment should rise in the severity of the crime (Becker, 1968) implies that men have a stronger incentive than do women to deter infidelity.

Households that allow husbands to deter female infidelity are ones in which the man contributes most of the market income. Infidelity is then deterred by husbands’ withholding consumption opportunities from them. Such deterrence is efficient if it brings about the desired outcome at less cost than alternative means. Withholding consumption opportunities appears to be at least a good candidate for efficient punishment. None of this is intended to stereotype women as being unable to survive without the help of a financially superior husband; the purely positive point is merely that pairs which allow husbands greater security of paternity will tend to be ones in which men out-earn their wives, so that wives get more by remaining faithful.

This paper constructs a model of household formation around these biological incentives. The model also predicts, inter alia, that the income gap will decrease in relative terms as families’ incomes increase. This is essentially because husbands must be in a position to deter disloyalty in their wives in order to avoid spending scarce resources on raising children that are not theirs. The relative income gap decreases in total market income because a stream of income which proportionately improves upon the prospects of an unmarried woman (or a married but disloyal and discovered woman) by only a little bit is nevertheless a big absolute increase when incomes are high. The only non-biological explanation for why husbands should be particularly wary of contributing less than precisely fifty per cent of their households’ market income in the literature is due to Bertrand et al. (2015), and is based upon male identity demanding that a husband makes more market income than does his wife20.

The model also generates a number of other implications. For instance, higher male (permanent) income inequality implies that wives face a greater risk when committing

19 There are several treatments of population-wide outcomes of genetically encouraged behaviour which deal extensively with altruism within the family due to the genetic links of the members, including classic treatments such as Sociobiology by Edward O. Wilson, or popular ones like Richard Dawkins’ The Selfish Gene. Economists have also discussed evolutionary-biological bases for why fathers should care about their own children (e.g. Becker, 1976 and 1981; Posner, 1992), or why firms maximize profits in the market (Alchian, 1950).

20 There are of course many rationales for why men earn more than women (e.g. discrimination, childbearing, or socially-conditioned preference differences), but they do not imply that market income differences should be legion also within households. Within-household specialization similarly fails to account for the significance of the fifty per cent mark in relative contributions to household market income by husbands.
adultery, since, in case their affair results in their husband’s divorcing them, new matches are less likely to be with men of a standing similar to the one of the most recent husband. Being associated with greater risk, adultery committed by women should therefore decrease the more male (permanent) income inequality there is. Consequently, husbands in high-inequality societies must somewhat paradoxically out-earn their wives by a smaller magnitude than must husbands in low-inequality societies, ceteris paribus. For the same reason, the present model predicts a higher frequency of female homemakers in countries with lower rather than higher (permanent) income inequality.

Notice that these biological incentives are not meant to compete with other theories of household formation, but rather to complement them, mainly in the form of a constraint. For instance, the maximization of a joint product could take place under the constraint that the husband must out-earn the wife. The main contributions of the present article are therefore twofold: (1) the finding that men will out-earn their wives purely due to the biological fact that only women bear children and the biological hypothesis that men have a taste for raising their own offspring; and (2) the conditions under which men may have to out-earn their wives by even more, including implications for marriages from institutions and contraceptives, as well as from the aforementioned issues of wealth and inequality.

The remainder of this paper is structured as follows: Section II provides an overview of some of the related literature. Section III outlines a continuous-time model, following the practice in principal-agent (and search-and-matching) models seen in approaches such as Shapiro and Stiglitz (1984), Rocheteau (2001), or Mortensen and Pissarides (1999). This model captures the biological incentives outlined in this section, and the main results of husbands out-earning wives with the proportional difference decreasing in total household market income. Section III highlights some of the important predictions concerning male-female income differences and implications from inequality, as well as a number of additional predictions and discussion, contrasting the present article with other work in family and labour economics. Section V summarizes and concludes. Throughout, the discussion will deal with ‘husbands’ and ‘wives’ but it should be obvious that unmarried heterosexual couples in deep relationships are equally part of the analysis.
II - BACKGROUND

Incentives from biology have been incorporated in other treatments’ explanations of various aspects of household behaviour. For example, the high incomes of prostitutes have been adduced to the importance of commitment to husbands. Prostitutes require large compensating differentials for their work since they cannot land a husband by offering sexual exclusivity (Edlund and Korn, 2002). On the same theme, Rai and Sengupta (2013) analyse pre-marital confinement of girls and young women such as seclusion, foot binding, and other practices to induce or enforce docility as signalling devices of future fidelity.

Saint-Paul (2015) provides a model in which men pay for the legitimacy of their progeny by transferring consumption possibilities to their wives, which leads to hypergamy (women marrying men of higher socioeconomic standing than themselves) since high-income men have low marginal utility of consumption, whereas the opposite is true of low-income women. Since women who marry average men forego mating opportunities with high-quality males, the result is that men marry down and women marry up. Thus, the highest-quality women have trouble finding husbands.

There is rather strong evidence for hypergamy. Historical practices such as bride prices and the absence of groom prices underscore the tendency of husbands’ (or their families’) out-earning wives (Edlund, 2013, pp. 654-656 provides a good discussion). Studies on courtship spending tend to be similarly biologically motivated and focused on male expenditures (Bergstrom and Bagnoli, 1993; Bronsert et al., 2017). In modern times, marriage rates rise more strongly with income for men than for women (CPS data reported in Saint-Paul, 2015), and the aforementioned study by Bertrand et al. (2015) shows striking distributions in which the frequency of marriage drops sharply just after the point at which men earn fifty per cent or more of household income.

As these papers indicate, sexual and marital mores which stress the fidelity of wives and the economic superiority of the husband are ubiquitous. This pervasiveness suggests that social norms are less simple a foundation than is biology for understanding manifold, disparate practices of arguably rather similar effects (i.e. inducing female fidelity in one form or another). Since the technical aspects of procreation are the same everywhere, sexual and marital mores of similar effects are apt to be the consequences of biological forces, and the premiss follows that the male disadvantage in ascertaining paternity implies higher costs for males of female adultery.

What are the incentives for women to commit adultery? From the biological point of view, a wife may be interested in cuckoldi ng her husband because doing so diversifies the genetic material with which she mixes hers and thereby enhances the viability of her progeny. In addition, there is the “sexy son”-hypothesis (Fisher, 1930), which says that infidelity with a ladies’ man (i.e. a man possessing characteristics prized by many women) that results in a son is apt to make him, like his father, an unusually successful procreator, which in the long run spreads the wife’s genes farther. Of course, many factors that might not be biological will likewise incentivize female (as well as male) infidelity, such as romance, the dullness of one’s husband, or the excitement of a paramour.

There are theoretical exceptions to the view that men respond negatively to cuckoldry. For instance, Liedtke and Fromhage (2012) develop a model in which (half-)siblings care for one another, so that if the male householder is the father of but some fraction of his wife’s children, his caring for all of them is still optimal because the (half-)siblings all help one another. The husband’s offspring will benefit from fitter half-siblings, so the family survives.

However, virtually all thorough empirical studies indicate that non-paternity is rare, with estimates ranging from less than one per cent of children in a German study using
bone-marrow transplantation samples (Wolf et al., 2012), to between one or two per cent per generation in a multi-century Flemish study based on a combination of genealogical and surname data (Larmuseau et al., 2013). Studies with a greater focus on groups of lower socio-economic standing tend to produce somewhat higher estimates, occasionally in the area of ten per cent or more (an example is Cerda-Flores et al., 1999, though they do not explain their sample selection in any great detail).

Longitudinal studies on cuckoldry fail to show large changes over time even though they span periods over which modern contraceptive aids became increasingly available. That the incidence of non-paternity is almost uniformly very low but does rise on occasion is an indication that husbands by nature choose rather good ways of deterring adultery by their wives, adapting to different technologies and institutions as necessary. Plausibly the vagaries of income trajectories under the circumstances when this deterrence fails make it more difficult to maintain relative dominance in household market income. Additionally, other men may be more likely to be of better quality when a wife is married to a very low-status one.

The focus on deterrence of infidelity in the present approach differs from those taken by the related biologically-motivated literature and produces an array of empirical implications. For instance, since the fifty-per-cent mark in husbands’ contributions to household income is an outcome of special significance in the present paper, there is no reason to expect the highest-earning women to remain unmarried (as in Saint-Paul, 2015) because they only have to earn less than do the highest-earning men. Indeed, the proportion of never-married women aged 35-39, as reported in Saint-Paul (2015), is fairly flat across incomes; 15.7 per cent of women earning less than $5,000, 18.1 per cent of women earning more than $100,000. The corresponding figures for men are 44.4 and 9.6 per cent, respectively. These patterns are consistent with most women marrying, and with men finding it increasingly difficult to sustain a marriage the lower their incomes fall.

By how much husbands wish to out-earn their wives is also affected by institutional and technological factors, variations in which allow adultery to be committed by wives with different probabilities of impunity. If it is likely that a marital transgression will be discovered, men do not have to out-earn their wives by as much as when this is less likely. For instance, DNA tests for paternity increase the probability that an unfaithful wife be found out and consequently allow husbands’ incomes to be lower than otherwise.

Other predictions by the present paper are also consistent with a number of broad and widely-noted empirical tendencies. If husbands out-earn their wives and most people marry, a gender ‘pay gap’ will spontaneously arise which, ceteris paribus, is greater in poorer societies and smaller in wealthier ones. Other explanations for the ‘pay gap’, such as lower female flexibility regarding work hours or different negotiating tactics for jobs, could be thought of as consequences of intra-household pressures on men to earn most market income (see Goldin, 2014, for an overview of potential explanations for the ‘pay gap’). Also consistent with the present paper are the oft-observed male marriage premium and female marriage penalty21.

21These names refer to the common findings that married men out-earn single men whereas single women out-earn married women. The male marriage premium is usually estimated to be around or above ten per cent (e.g. Bardasi and Taylor, 2008; Ginther and Zavodny, 2001), and have been attributed to one or several of ability bias (only productive men get married), signalling (marriage is certifies goodness), or human capital (marriage makes men more productive, maybe due to household specialization). The female marriage penalty is often estimated at around ten per cent in studies relying on the National Longitudinal Survey of Youth (e.g. Light, 2004). In contrast to the male premium, the lower wages of married women is more difficult to attribute to (in)ability bias. Specialization within the household (or employer perceptions that this will be the case, i.e. signalling) has usually been thought to be behind the reduction in female wages upon marriage.
That males contribute most of the market income within households has recently been suggested to be due to perceptions of male and female identity, *inter alia* because if the wife’s earnings potential is larger than that of her husband she is nevertheless more likely to be a homemaker or earn less than her husband if she does work (Bertrand *et al.*, 2015). This finding is hard to reconcile with explanations based on human capital or specialization; comparative advantages predict that in those cases where the woman’s earnings capacity exceeds that of her husband, she should account for most of her household’s market income and her husband should do more non-market work (Becker, 1973, p. 828). By contrast, the biological incentives suggested by this paper predict no such thing, since the husband must out-earn the wife for the marriage to work.

A focus on biological incentives can also turn apparently Pareto-suboptimal outcomes into efficient ones. For instance, Udry’s (1996) finding that male-controlled plots of land in Burkina Faso are cultivated more intensively than are female-controlled ones within the same family (so that their marginal returns are not the same) makes sense from the constitutional-choice perspective, since in the possibility of hidden action the husband must have resources sufficient to deter infidelity.
III – THE MODEL

In terms of structure, the model may be thought of as a principal-agent problem in which husbands are the principals who wish to rear offspring that share their genetic material. Wives are the agents who may work for them (produce a child with their husband), or shirk (produce a child with someone else). A husband uses part of his income to supplement that of his wife, but only so long as she remains faithful to him. Through divorce or simply through retraction, it is the prospective loss of this supplement that deters female adultery. In this sense, the model’s mechanism is reminiscent of the efficiency-wage model by Shapiro and Stiglitz (1984). Males wish to induce fidelity and females wish to maximize income, consisting of her own income and a part of her husband’s income.

Although wives surely dislike their husbands’ cheating on them, they cannot deter this behaviour as easily as can their husbands since biological differences imply that their ability to do so would threaten the viability of the household. Essentially, the biological perspective implies that women dislike spousal unfaithfulness less intensely than to men. It is an unfortunate fact for women that their husbands can commit adultery with impunity (from the perspective of household finances). When at least some people are unmarried sometimes, one extramarital affair need not imply another extramarital affair, which ensures that these household rules to deter adultery do indeed aggregate (cf. Masters, 2008, where this does not happen as easily because both husband and wife are committed).

The following establishes the essential building blocks of the simple model. Let $V_p$ denote the net present value for a married but unfaithful woman whose husband has diverted consumption possibilities from her (again, not necessarily by divorce), and consequently, until her husband relents or until she is remarried, has to rely on only her own income. In other words, $V_p$ is the value function of a woman whose infidelity has been discovered. The woman’s own income is denoted by $(1 - h)w_{Hht}$, where $w_{Hht}$ is total household market income, and $h$ is the husband’s contribution to it (more on $V_p$ below). $V_M$, $s$, and $r$ denote the value function of the married but unfaithful wife, the share of total household market income consumed by the wife, and time preference (or the rate of interest), respectively.

The probability of being found out when unfaithful is treated as exogenous and is denoted by $p$. The possibility of institutional features to impact $p$ is discussed in greater detail below. The punishment comes with a varying degree of permanence, captured by a parameter $\rho$, which is below one in cases where adulteresses are considered pariah or – in case of divorce – has poor marriage-market prospects, and above one for wives whose marriages are simply bad matches (which encourages adultery and eases its consequences). A related but distinct variable, $m$, captures a punishment rescindment rate (though it may equally be thought of as a remarriage rate). For example, with $\rho$ substantially less than one and $m$ close to one, disloyal and divorced women would get remarried fairly easily, but only to men of substantially lower calibre than their former husbands (high punishment permanence).

A wife can choose between fidelity and infidelity. Looking ahead a short period of time $[0, \Delta t]$, the stream of payoffs for infidelity is written

$$
V^U_M = sw_{Hht}\Delta t + e^{-r\Delta t}[p\Delta tV_p + (1 - p\Delta t)V^U_M].
$$

That is, at every instant, she receives a share $s$ of the intensity of her household’s pooled market income at each moment in time and, with appropriate discounting, continues to do so unless found out, in which case she gets $V_p$. The probability $p$ of being found out is
constant per time unit. Since \( e^{-r\Delta t} \approx (1 - r\Delta t) \) for small values of \( \Delta \), the above expression turns into

$$V^U_M = sw_{HH}\Delta t + (1 - r\Delta t)[p\Delta t V_p + (1 - p\Delta t) V^U_M].$$

Isolating \( V^U_M \) on one side yields

$$(p\Delta t + r\Delta t - \Delta t^2 rp)V^U_M = sw_{HH}\Delta t + (1 - r\Delta t)p\Delta t V_p.$$  

The above expression can be rearranged into

$$V^U_M = \frac{sw_{HH}\Delta t}{p\Delta t + r\Delta t - \Delta t^2 rp} + \frac{p\Delta t V_p}{p\Delta t + r\Delta t - \Delta t^2 rp} - \frac{\Delta t^2 rp V_p}{p\Delta t + r\Delta t - \Delta t^2 rp}.$$  

Eliminate duplicate \( \Delta t \)'s from numerators and denominators:

$$V^U_M = \frac{sw_{HH}}{p + r - \Delta trp} + \frac{p V_p}{p + r - \Delta trp} - \frac{\Delta trp V_p}{p + r - \Delta trp}.$$  

As is common practice in continuous time, the growth rate of this value function is obtained by taking limits as \( \Delta t \rightarrow 0 \). When \( \Delta t \) gets smaller, the growth rate approaches

$$V^U_M = \frac{sw_{HH}}{p + r} + \frac{p V_p}{p + r},$$

Multiplying by \( (p + r) \) this expression yields equation (1):

$$rV^U_M = sw_{HH} + p(V_p - V^U_M).$$

Faithful wives, like unfaithful ones, get a share \( s \) of the household’s total market income \( w_{HH} \) and lose the opportunity cost \( c \) of foregone extramarital mating opportunities (in lieu of this opportunity cost, one could posit a benefit to having an affair in equation (1), which comes to the same thing). This happens at all times \( t \) and goes on into the indefinite future. The interest rate is again \( r \).

$$V^F_M = (sw_{HH} - c)\Delta t + e^{-r\Delta t}V^F_M.$$  

We look ahead a short period of time \([0, \Delta t] \). Since \( e^{-r\Delta t} \approx (1 - r\Delta t) \) for small values of \( \Delta \), the above expression is

$$V^F_M = (sw_{HH} - c)\Delta t + (1 - r\Delta t)V^F_M,$$

which may be rewritten as

$$r\Delta t V^F_M = (sw_{HH} - c)\Delta t.$$  

Divide both sides by \( \Delta t \) to obtain equation (2)

$$rV^F_M = sw_{HH} - c.$$  

---

\( ^{22} \) The present model could alternatively be set up in discrete time with a value function corresponding to (1) above of the form \( V^U_M = sw_{HH} + \delta (pV_p - (1 - p)V^U_M) \), where \( V_p = (1 - h)w_{HH} + \delta (1 - h)w_{HH} + \delta^2 (1 - h)w_{HH} + \ldots + \delta^{r-1} (1 - h)w_{HH} + \delta^r V^U_M + \delta^{r-1} V^F_M \) and \( V^F_M = sw_{HH} - c + \delta V^F_M \) for unfaithfulness and faithfulness, respectively.
Solving (1) and (2) for \( V_M^U \) and \( V_M^F \) yields

\[
V_M^U = \frac{SW_{HH} + pV_p}{p + r},
\]

and

\[
V_M^F = \frac{SW_{HH} - c}{r}.
\]

Women are faithful when \( V_M^F \geq V_M^U \), i.e. when \( \frac{SW_{HH} - c}{r} \geq \frac{SW_{HH} + pV_p}{p + r} \), or equivalently the condition that

\[
sw_{HH} \geq rV_p + \frac{(r + p)c}{p}.
\]

This expression says that the share of household market income that goes to the wife must not fall below her stream of benefits if she loses her husband’s income supplement, plus a fidelity bonus that increases in the opportunity cost of fidelity and in the discount rate, and falls in the probability that a wife’s unfaithfulness is discovered.

Women who have been unfaithful and discovered receive \( rV_p = (1 - h)w_{HH} + m(\rho V_M - V_p) \), where \( h \) is the husband’s contribution to household income, \( V_M \) is the married woman’s expected utility, and \( m \), as mentioned above, is the per unit of time degree to which her transgression is “forgiven”. Equally, \( m \) may be thought of as the punishment rescindment rate or the remarriage rate, if divorced. The term \( \rho \) (as also discussed above) captures a permanent part of the punishment (for \( \rho < 1 \), as divorcées have at times been commonly held in disrepute. The reverse case, \( \rho > 1 \), may be interpreted as the present match being a bad one, which encourages adultery.

Stable marriages are clearly characterized by fidelity, so that \( V_M = V_M^F \). Using this condition, one may insert equation (4) into the expression for \( rV_p \) to obtain

\[
rV_p = \frac{w_{HH}(r - rh + mps) - mpc}{r + m}.
\]

Equation (6) says that a disloyal and discovered woman is unambiguously better off the higher is her household’s income, and more so the lower is her husband’s share of household income, but she would be better off regardless of \( h \). This is partly because whatever her husband’s share, a larger household income means that her share represents more absolute income. In addition, the possibility of punishment rescindment (for high household incomes relative to \( c \)) offers additional security\(^{24}\). The greater the part of her punishment is permanent (i.e. the lower is \( \rho \)), the worse off she is provided that her share of household income is greater than the opportunity cost of fidelity.

There is an alternative interpretation of the parameter \( \rho \). Rather than capturing the permanence of a punishment, it may equally be thought of as saying something about the likelihood that a divorced woman will find a similar-calibre husband for her next try

\(^{23}\) This comes from \( V_p = mV_M + (1 - m)V_p = mV_M - mV_p + V_p = m(V_M - V_p) + V_p \), where \( V_p = (1 - h)V_M \).

\(^{24}\) Differentiating (6) with respect to \( m \) yields \( \frac{(Sw_{HH} - c)(r + m) - w_{HH}(r - rh + 0.5m) - mc}{(r + m)^2} \), which is positive for \( h \geq 1 + \frac{c + r + 2cm}{w_{HH}r} \), i.e. when the husband earns at least half of the household’s income and the household’s income is sufficiently greater than \( c \).
at matrimony. This interpretation carries some predictions regarding the role of male (permanent) income inequality because if the chances of finding a similar man go down then so does $rV_p$, implying fewer instances of adultery in societies with more (permanent) income inequality among men, ceteris paribus. Because female adultery is associated with greater downside risk, such higher-inequality societies also have less pressure on males to out-earn their wives.

Substituting (6) into (5) and rearranging yields the fidelity condition on the husband’s share of household income

$$h^* \geq \frac{c}{w_{HH}} \left( \frac{p + r + m}{p} + \frac{m - m\rho}{r} \right) + 1 - s \left( 1 + \frac{m - m\rho}{r} \right).$$

Since deterrence means that the wife’s contribution to household income cannot exceed her consumption of it, a lower $s$ requires a higher $h$. To see the effect of greater opprobrium for unfaithful women (i.e. when $\rho$ declines), differentiate (7) with respect to $\rho$; husbands must contribute more to household income if the wife’s consumed share of household income exceeds her opportunity cost of fidelity, but the husband may contribute less if $sw_{HH} < c$. That is, a more severe permanent punishment makes a high opportunity cost of fidelity more bearable (it is not worth the risk) and $h$ may fall, but if $c$ is negligible to begin with, $rV_p$ in equation (6) is greater, meaning that unfaithful and discovered women are better off and husbands must compensate by a higher $h$. Under the interpretation of $\rho$ as the probability of finding a new, similar-stature husband for a divorced woman, the above result on male (permanent) income inequality is supported. These derivations lead to the following proposition.

Assume now that there is neither permanent punishment nor any terrible matches so that $\rho = 1$, and assume that households share resources equally so that $s = \frac{1}{2}$. Under such circumstances, the following result holds.

RESULT: The husband’s contribution to his household’s total market income must exceed fifty percent to incentivize his wife to remain loyal to him and not have an extramarital affair, but may fall arbitrarily close to this mark as total household market income rises, assuming that husband and wife divide total household market income equally between them, so that the wife gets one half in monogamous marriages.

PROOF: Rewrite equation (7) as

$$h^* \equiv h \geq \frac{1}{2} + \frac{c}{w_{HH}} \times \frac{m + r + p}{p}.$$

It is evident from equation (8) that a division of household market income other than fifty-fifty would establish other thresholds, but in monogamous marriages fifty-fifty (or something very close to it) would seem like the natural division. ■
IV – IMPLICATIONS

The foregoing model does not deal with certain factors which will plausibly impact income as well as opportunities for extramarital affairs. Working longer hours is likely the best example of one such factor, since a husband can thereby increase his relative income within his household, but also loses time he could have spent with his wife. The consequent lack of companionship may be an inducement to infidelity.

The model abstracts from this potentially counteracting force because the probability of discovering infidelity depends on many other factors than how much a husband works when his wife does not. Friends, family, or acquaintances may run across a person’s wife and casually mention this to her husband, resulting in the wife’s having to make up stories to explain her whereabouts. Suspicions may also be aroused from differences in behaviour. More fundamentally, with significant assortative matching, men and women marry who are at the same rank within their sexes. Thus, a man would not typically have to be at work much more than his wife in order to out-earn her.

Since the model predicts that husbands must out-earn wives by more the lower are their joint household incomes, time at work is most likely to counteract deterrence of infidelity in very poor societies. Therefore, it would be consistent with the present model if women had more freedom to move around the richer are their societies. But again, this holds only if husbands’ time spent at work is a significant contributor to marital transgressions by women.

Turning now to some of the main implications of the foregoing section, it is evident from equation (8) that the more impatient the wife is (i.e. the higher is \( r \)), the higher \( h \) must be. Furthermore, it is easy to see that the effect on the critical share of the husband’s contribution \( h^* \) of an increase in the probability that an adulteress is found out, \( p \), is to make it go down, which is in line with the classic findings in law and economics that one need not be able to punish so severely if a crime is easily detected (Becker, 1968). An increase in the social acceptability (or punishment rescindment rate, or remarriage rate) of marital infidelity, \( m \), puts upward pressure on \( h^* \), in line with intuition: for whatever punishment there is, if it is felt for a shorter time it must be felt more greatly in order to deter wrongdoing.

The effect of an increase in total household income is to drive \( h^* \) asymptotically down towards one half. This captures some of the long-run developments that have occurred in many advanced countries, where the point at which the husband’s income just exceeds fifty per cent has become increasingly common among married couples (Bertrand et al., 2015, p. 579). The outcome illustrates how a slight excess over fifty per cent will represent absolutely larger consumption possibilities usable for punishment the higher is household income. This point also draws more mass as household incomes rise because a relatively small but absolutely large difference induces “punished” women to seek remarriage or rescindment (\( m \) is positive).

Indeed, a striking finding in family economics in recent years is that of the sharp, discontinuous decline in the distribution of husbands’ contributions to their households’ income just after the point at which they earn fifty per cent. The discoverers, Bertrand et al. (2015), attribute this finding to male identities being inconsistent with contributing less than one half. The present model also finds a special significance of this point in relative contributions to household income, and thereby offers an alternative explanation to that of identity.

Although an established and important part of economic science, identity-based theories of household behaviour face the hurdle of explaining why perceptions of identity remain so stable over time even amidst great financial incentives for them to change. As Bertrand et al. (2015) note, wives fail to specialize in market work even when their
predicted earnings capacity exceeds that of their husbands. Biological incentives and identity economics are not mutually exclusive explanations, so biology may help to account for the money left on the table.

One recent finding that speaks against identity-based explanations is due to Zinovyeva and Tverdostup (2018), who look at relative incomes within households for Finnish couples. They also find a sharp discontinuity at fifty per cent, but when slicing the data more thinly they are able to attribute all of it to couples who work closely together – mainly in the same (small) firm, but extending to the same industry. The reason is that the point at which husbands and wives have roughly equal incomes has more mass while otherwise husbands continue to out-earn their wives (as in couples not working together). This finding is consistent with the present model, since working closely together implies that the probability that infidelity is found out will be significantly increased, which eases the pressure on a husband to out-earn his wife.

In equation (7), a higher \( \rho \) raises the husband’s required contribution \( h \) when the share of household market income consumed by the wife is greater than her benefit of an extramarital affair (as will be the case when husbands successfully deter adultery by their wives):

\[
\frac{\partial h}{\partial \rho} = -\frac{cm}{w_{HH}r} + \frac{sm}{r} = -c + s_{HH},
\]

which is greater than zero when \( s_{HH} \geq c \). In other words, if finding a husband of a similar income to one’s most recent one becomes likelier (i.e. \( \rho \) goes up), the husband must contribute a greater share of the household’s total market income. Since such new husbands are harder to find the more unequal a society is, one would expect such societies to be characterized by (1) more equal relative contributions to total household market income by husband and wife, and by (2) less adulterous behaviour on the part of wives.

Such tendencies hold if all else is held equal. Under this interpretation of \( \rho \), it would also seem natural to think that a divorced woman’s finding a husband of similar standing to her previous one should be still more difficult at the top, where the number of men with higher incomes than the previous husband is necessarily lower than at the bottom of the income distribution (given the usual income distribution in which the median income is far below the mean). Additional implications follow.

**Income Trajectories and Labour-force Participation:** Since (nearly) equal contributions to household market income are more likely for higher total household income, wives’ wages are expected to be suppressed more at the bottom than at the top of the income distribution. It follows that increasing married female labour-force participation occurs first among relative high-income earners, and last among low-income earners, while wives’ incomes follow a similar trajectory of being closest to those of their husbands at the top and taking longer to achieve near-parity the worse-off are husbands at the bottom.

Such pressures, in turn, encourage assortative matching, because women who can earn a relatively high market income have incentives to seek out similarly capable men, whose relatively high position in the income distribution makes them least likely to object to having a high-earning wife\(^{25}\). At the very bottom where pressures to make money may be overwhelming, increases in female labour-force participation are apt to

\(^{25}\) Note that assortative matching relates to individual ranks in gender-specific income distributions. While models of assortative matching traditionally place no importance on the relative incomes of husband and wife, introducing fear of cuckoldry tips the balance to favour men as the primary income earners, with women approaching the fifty per-cent mark as incomes rise higher.
come at the price of familial disharmony as men substitute other means to deter infidelity for a high income, and of increasing rates of non-paternity (as in the evidence discussed in Section II).

For groups of low socioeconomic standing, women are also presumably paired with men of below-average quality, which means that opportunities for extramarital affairs are likelier than otherwise to be with men of higher quality than their husband, when societies are sufficiently heterogeneous that higher-quality males are around. This raises $c$ in the present model and strains relationships. This would also be consistent with Bertrand et al. (2015)’s finding that the discontinuity at the husbands’ side of the fifty per cent mark of household income contributions is thicker for poorly-educated couples than for highly-educated couples (online appendix, p. 25).

Specialization: Biological incentives share with human capital considerations the implication that husbands will specialize in market work while wives specialize in household work, giving rise not only to the male marriage premium, but also to the female marriage penalty (Becker, 1991, pp. 37-48). In contrast to human capital theory, biological incentives suggest that same-sex households have less pressure than others to specialize in household and market work since any progeny cannot be related to both householders. Studies on same-sex households have produced mixed results on specialization, but generally find that there is less of it among same-sex cohabitants of both homo- and heterosexual inclinations (see Zavodny, 2008, for a thorough overview and study, which also finds no cohabitation premium for same-sex individuals irrespective of sexual preference).

Another addition to explanations based upon human capital is that infertile women and their husbands should have similar labour-market outcomes as fertile married couples. Although the husband in such couples knows that his wife cannot give birth to another man’s child, he should nevertheless desire to contribute the majority of total household income because attraction and pair-bonding have a biological basis and it is unlikely that we should have evolved preferences over household contributions which depend upon such inconspicuous details of the mate’s characteristics.

Institutions: Theories based purely on human capital and specialization have difficulties explaining why wives earning most of a household’s market income may do more household work than do their husbands. The finding that this is what happens is due to Bertrand et al, (2015), although it has not been replicated by German data (Wieber and Holst, 2015). Bertrand and her co-authors argue that women compensate for having incomes that conflict with the male identity of being the main contributor to the household. However, wives’ accounting for most market and non-market work might also be seen as a seclusion mechanism which guards husbands from the threat of cuckoldry by limiting the wife’s opportunities for extramarital affairs. This institutional characteristic would have the effect of raising $p$.

This mechanism is akin to how, at various times and in different places, many wives have been secluded in the more literal sense of being put in harems, as well as through mechanisms such as foot binding and other practices mentioned in the introduction. Raising $p$ by limiting the wife’s time essentially does the same job as do more tangible acts of legislation or social norms in reinforcing the censure of female adultery. Seclusion mechanisms presumably reduce the woman’s earnings capacity, thereby enabling low-earning men to marry a woman (and higher-earning men to marry several women) knowing that a husband’s withholding of consumption opportunities from his wife (wives) will be an effective deterrent of infidelity.

In line with commentary on unintended effects of welfare programmes (e.g. Murray, 1984), increased payments and greater availability of social assistance would put greater
strain on low-income marriages, since outside options improve for wives of low-income men. Examining the impact of the Welfare Reform Act in the US of the 1990’s, in which a more stringent assistance system was implemented, Bitler et al. (2004) estimate that divorce rates did indeed fall upon Welfare Reform. Lemieux and Milligan (2008) find more suggestive evidence of the same phenomenon in a regression-discontinuity study of welfare reform in Quebec.

Similarly improved outside options can come from prenuptial agreements, joint-access bank accounts, and the vagaries of divorce settlements, since they make it more difficult for the husband to divert consumption possibilities from the wife. These issues immunize wives at the higher end of the income distribution from the threat of diversion of consumption possibilities by their husbands, thereby lowering the pressure on men to earn at least half of their household’s market income, and imperilling the fidelity of their wives. Wives may here have an incentive to pre-commit by limiting their access to joint bank accounts, or not asking for prenuptial agreements (which are quite rare). It may be harder to pre-commit to an unfavourable divorce settlement, though oftentimes perceptions among the general public are such that they give the majority of the estate to the husband (Hersch and Shinall, 2017).

**Technology:** That husbands have always been loath to lose exclusive sexual access to their wives is a traditional belief in family economics. In discussing adultery and its effects on marital stability, Becker (1991, p. 48) notes that “men are reluctant to rear children fathered by others”. Similar reasoning is found in Posner’s (1992) discussion on why polyandry (one woman’s marrying more than one man) is so rare. Advances in contraception and paternity testing allow husbands to rule out non-paternity, which may seem to eliminate the harm done by an unfaithful wife. However, because the preferences for female fidelity are rooted in biology, it would take a great many generations for recent technological advances to have any effect on genetic selection of preferences that do not mind female adultery to become a substantial fraction of the male population. Hence, these advances affect \( p \) (and thereby \( h \)) rather than preferences.

Contraceptives may reduce the probability that adulterous wives are discovered by eliminating physical evidence of extramarital affairs (less risk of conception). To this extent, contraceptives put upward pressure on husbands’ wages relative to those of their wives since the low probability of detection requires a sterner punishment. However, since contraceptives reduce the risk of extramarital affairs, they may be engaged in with greater frequency, which implies a higher probability of discovery and downward pressure on husbands’ wages. DNA-based paternity tests, on the other hand, unambiguously increase \( p \).

**Perceptions and Preferences:** Lastly, the present model is also consistent with the empirics of male and female expectations of, and preferences for, characteristics in males and females. Hersch and Shinall (2017) devise a survey with a hypothetical divorce case and varying occupational backgrounds of a husband and a stay-at-home wife, whose previous career either matched that of her husband or earned her a lower income.

---

26 This is not strictly true; paternity tests can yield erroneous results. Michael (1974) estimates the incidence of conception when a couple has regular intercourse using a variety of contraceptive techniques. For even a 99-per-cent effective contraceptive, the probability of conception still exceeds ten per cent in a five-year interval (p. 133).

27 Paternity tests are not uncommon in the US; there were around 280,000 DNA-based paternity tests performed annually in the early 2000’s (Bishai et al., 2006) out of nearly four million births in 2014 (of which approximately 40 per cent were to unmarried women) reported by the CDC: https://www.cdc.gov/nchs/fastats/births.htm. The number of tests relative to the number of births overstates the incidence of doubtful cases, however, since one child may have several possible candidate fathers.
Whether male or female, the respondents invariably award the husband the majority of the household’s assets. Though the authors do not state it, these results are consistent with a common expectation among respondents that the husband be household’s primary earner, so that he consequently may receive the majority of the household’s assets upon divorce.

In a study on speed dating, Fisman et al. (2006) present findings on desirable mate characteristics. From the present model, one would expect males to wish to marry women whose (permanent) market incomes are lower than are their own, and women to have the reverse preference, and in their discussion of prior literature, they report that this is essentially what psychologists have long tended to find (pp. 675-676). For instance, a partner’s earning potential is valued more by women than by men.

The paper by Fisman and coauthors is noteworthy because the authors are able to record several personality traits of their subjects through the speed-dating setup. Thus, they find that variables such as ambition, intelligence, and socioeconomic background are all valued more by women than they are by men. Ambition and intelligence are in fact disvalued by a man when a woman exhibits more of them than he does, which is precisely what is to be expected if the optimal marital rules require males to contribute more than half of household income.
V – CONCLUSION

This paper introduces a mechanism based upon biology that leads husbands to wish to out-earn their wives in order to reduce the latter’s incentives to commit adultery. From a biological perspective, adultery by wives harms husbands more than adultery by husbands harms wives, since only husbands may come to spend scarce resources on raising genetically alien material. Thus, the mechanism introduced by the present paper essentially rests on preferences which evolutionary biology suggests should be widespread.

The present approach in which husbands out-earn wives to deter infidelity seems to be consistent with many disparate empirical findings in labour economics, such as the male marriage premium and female marriage penalty, the gender “pay gap”, the widespread tendency of husbands to earn more than fifty per cent of their households’ market income, differences between the sexes in mate selection, and a reduced pressure for household specialization among same-sex cohabitants.

In addition, new implication from the present model include that of male (permanent) income inequality being correlated with a reduced pressure on husbands to out-earn their wives, some dynamics of female labour-force participation, as well as pressures on relative market incomes within households from technological and institutional developments. For traditional marriages more generally, the implication from the present model is that husbands will continue to tend to earn at least fifty per cent of household income, and that marriages which do not obey this rule are more susceptible to discord.
REFERENCES


EKONOMI OCH SAMHÄLLE
Skrifter utgivna vid Svenska handelshögskolan

ECONOMICS AND SOCIETY
Publications of the Hanken School of Economics


In this dissertation, four essays examine various ways in which people make decisions based upon others’ characteristics or actions. Individuals certainly can choose without regard to others and many decisions essentially lack social influence, such as the choice of one’s tie or scarf or car, but oftentimes peers may react to a choice in a way which decision-makers prefer not to disregard. Thus, some smoke or vote for the Greens because others do, or wear a certain haircut because it is the style at the time. Sometimes decisions may be private within some bounds, such as choosing any uni-coloured car, but not a hippie-style multi-coloured one.

This dissertation seeks to demonstrate the influence of such social elements on choices pertaining to four distinct areas: (1) whether to seek public assistance (the dole); (2) whether to vote or abstain; (3) what to eat when others may observe one’s choice; and (4) whom to marry to avoid some causes of marital disharmony.

The treatment is always applied and mainly empirical; only the fourth essay makes use of an exclusively theoretical argument. This essay also differs in reducing the scope of social influence to characteristics of only one’s spouse(s). The article suggests an empirical pattern in which husbands tend to out-earn their wives in order to be able to withhold income from them to reduce the probability of non-paternity, with differences between the sexes declining as total incomes rise.

The other three articles proceed in the opposite way, finding empirical patterns and suggesting explanations for them. The first essay finds that welfare offices in buildings with features that enhance the visibility of entry (such as a set of steps before the entrance), tend to approve a greater share of applications than do other welfare offices. The suggested explanation is that only the neediest individuals will accept the risk of being seen to be ‘welfare cases’, eliminating comparatively haphazard applications to raise the approval rate.

The second essay looks at data from countries practicing multiple concurrent national referenda, finding a pattern in which more concurrent referenda are associated with lower turnout, akin to quantity discounts being associated with fewer sales. This puzzling relationship, too, may be accounted for by peer effects. If groups are affected by referenda outcomes but individuals do not really wish to vote, group pressure may be less effective when several groups care about different referenda, because each group wants to avoid pressuring the non-voter.

The article analyzing dietary decisions finds that customers at a restaurant in Western Finland consume lighter meals when their body types are bigger than those of their peers. This finding points to a mechanism whereby thinness is prized and people wish to signal conformity to thinness-inducing habits by foregoing a larger meal in exchange for social recognition.