Essays on Experimental and Applied Economics

Submitted by

José Alberto Guerra Forero

for the degree of Doctor of Philosophy

of the

University College London

2014

Declaration

I, José Alberto Guerra Forero, confirm that the work presented in this thesis is my own. Where information has been derived from other sources, I confirm that this has been indicated in the thesis.

Signed.................................. (José Alberto Guerra Forero)
Date:

To my beloved grandparents, Alberto and Elisa, whose righteous example reverberates through all my actions. To Panela, mi único refugio posible, for awakening my mornings and banishing my nightly fears.

Abstract

In this dissertation I exploit observational and experimental data to study individual decision making when agents face social interactions or are described by non-standard neoclassical preferences.

In the first chapter I study how social interactions, could explain occupational choice in an incomplete information setting. In a discrete choice framework I allow for group unobservables affecting decisions. I show that asymmetries in the peer influence enables to separately identify the effects from group members' expected behaviour and the effects from their characteristics. I provide an empirical application to nineteenth century London. The results show that social networks were important in determining occupations but are somewhat lower than estimates which do not impose consistent beliefs nor allow for unobservables.

Secondly, I implement an artefactual field experiment with small entrepreneurs. Subjects were given an initial amount of money to be invested across alternatives. Some of the subjects were informed about the possibility of getting either a high or a low income level. The income level was either predetermined or allocated after a fair lottery. Agents who started with a low income after the lottery were more risk loving. A model of reference–dependent preferences with multiple reference points, formed through recently held expectations on foregone and actual outcomes, fits most of the experimental results.

In the last chapter I study game interactions in interdependent value auctions following Kim (2003). Agents are asymmetrically informed in terms of how precisely they know the different aspects of the object's value. Due to the mismatch of bidding strategies between these bidders, the second-price auction is inefficient. The English auction has an equilibrium in which bidders can infer information and attain efficiency. The increase in perfectly informed bidders increases the seller's revenue. A laboratory experiment confirms key predictions about efficiency and revenues and reveals naive bidding.

Acknowledgement

I am grateful to Syngjoo Choi for all his support. His close supervision and insightful teachings during these past few years made me a better economist and certainly a better person. Aureo de Paula's guidance in the later part of my studies substantially improved my work. The early discussions with Georg Weizsäcker of my research encouraged me to be more rigorous in my claims and left a lasting mark on the way I now approach new projects.

I thank Orazio Attanasio for allowing me to be part of the EDePo at the Institute for Fiscal Studies. My time spent there was an extremely stimulating academic experience. Thanks to my co-authors Jinwoo Kim and Myra Mohnen for sharing their insight with me.

I thank my parents and siblings, their continuous encouragement helped me pursue my aspirations. I thank my friends left behind in Colombia whose close company, no matter time and distance, was felt during these years of intense learning. I am indebted to my friends in London, for they taught me that national borders are now superfluous and should become obsolete.

The financial support from the Colombian Central Bank, IFS and COLFUTURO is appreciated.

Contents

List of Figures

List of Tables

Contribution to chapters

The chapter 2 is coauthored with Myra Mohnen (UCL).
The chapter 3 is a solo document.
The chapter 4 is coauthored with Syngjoo Choi (UCL) and Jinwoo Kim (Seoul National University)

Chapter 1

Introduction

In this dissertation I exploit observational and experimental data, together with microeconomic models, to study individual decision making when agents are subject to social interactions or are better described by non-standard neoclassical preferences. Social interactions are important for labour market or goods acquisition decisions when individuals are asymmetrically informed or have interdependent valuations. Non-standard preferences, in the shape of reference points against which one can compare outcomes from actions, could better describe investment decisions when individuals hold expectations about potential outcomes.

In chapter 2 I study how non–market institutions, mediated through local interactions among members of the same social group, could explain occupational choice in an incomplete information setting. This chapter presents a multinomial choice model with social interactions and heterogeneous beliefs. Individuals form heterogeneous expected behaviours of peers by taking into account their characteristics and the strength of their ties. Given the asymmetries on peer influences, even when unobservables hit the group as a whole, the effect from group members' expected behaviour (*endogenous effect*) and from peers' characteristics (*contextual effect*) are separately identified. I provide an empirical application to nineteenth century London and explore the importance of social networks in determining occupational choice. As ecclesiastical parishes were at the heart of social identity, groups are delimited by their boundaries. Using a novel dataset that pins residential locations down to the street level, I measure the strength of ties between members of a group based on geographical proximity. The unique two-tier administration which attributed the public good provision responsibility to a grouping of parishes allows me to mitigate the self-selection bias. My results show that social networks were important in determining occupational choice. Failing to allow for correlated effects leads to the overestimation of the true endogenous effect. Once multiple equilibria and group unobservables are accounted for we still find significant and positive effects for individuals unemployed and in industrial occupations, while a significant and negative effect for commercial occupations. Social interactions do not seem to matter for domestic and professional occupations.

Then, I investigate how reference–dependent preferences affect the individual decision making process and thus can help to understand economic phenomena. Such reference–dependence is characterized by preferences that have embedded points against which individuals weight the outcomes from their choices. Falling behind them is felt as a loss that looms larger than a commensurable gain when surpassing them.

My research aim in this area is to understand the nature of reference points, in particular whether expectations could determine them. In chapter 3 I implement an artefactual field experiment on a sample of Colombian informal small entrepreneurs. Subjects were given an initial amount of money to be allocated across different investment alternatives. A high or low initial amount of money was given to the subjects (*Income Variation*). Further, some of the subjects were informed about the possibility of getting either income level which was then to be determined by a coin toss, while the rest was assigned a predetermined income level (*Reference Point Variation*). The experimental results suggest that agents who started with a low income were more prone to invest if they faced the lottery on possible initial incomes compared to all other treatments. A model of reference–dependent preferences with multiple reference points is provided which fits most of the experimental results. In such model, recently held expectations on foregone and actual outcomes play a role in the decision making. Those with a predetermined initial income decide as if that income constitutes their reference point, while those who faced the initial lottery behave as if weighting their decisions against the foregone and actual one.

In the last chapter 4 I continue studying game theoretical interactions with asymmetric information, but now in an interdependent value auction setting. I follow Kim (2003)'s model were agents are asymmetrically informed in terms of how precisely they know about different aspects of the object's value and study two standard auction formats – the second-price (sealed-bid) auction and the English auction. Some bidders are perfectly informed about their value (*insiders*), while the rest only know the private component of their value (*outsiders*). Due to the mismatch of bidding strategies between informed and imperfectly–informed bidders, the second-price auction is inefficient. The English auction has an equilibrium in which imperfectly–informed bidders can infer, from the history of prices at which other bidders drop out, relevant information to attain efficiency. The increase in insider information by switching an outsider to an insider has a positive impact on the seller's revenue reminiscent of the linkage principle of Milgrom & Weber (1982*a*). A laboratory experiment confirms key theoretical predictions about efficiency and revenues. Nevertheless, there is evidence on naive bidding which declines in the number of insiders in the English auction.

The present thesis contributes to the literature on several fronts. First, chapter 2 adds to the scarce literature on multinomial choice models with social interaction by allowing for group–level unobservables and a network structure.

Secondly, due to the lack of information on actual contact, researchers usually proxy the relevant group using some arbitrary metric of distance based on social and/or geographical proximity (Gaviria & Raphael 2001, Hanushek, Kain, Markman & Rivkin 2003, Hoxby 2000, Sacerdote 2001, Solon, Page & Duncan 2000) In the modern world, there is a legitimate concern that physical distance may have become less and less important in shaping social networks. Therefore, focusing on this historical time period provides the advantage that I have a more credible proxy for social networks as interactions were mostly geographic in nature. Moreover, the religious property of my measure offers an additional dimension to social networks.

Third, thanks to the unique administration layout of the period studied, I provide plausible reasons why adding a fixed effects at the administrative level might better control for self-selection into a group.

The second chapter adds a piece of evidence on the role expectations play in shaping reference points. In particular, the experimental results suggest that recently held expectations on both: possible income levels (i.e. counterfactuals) and actual experienced income level affect individual decision making. Therefore, foregone opportunities as well as actual experiences serve as benchmarks against which individuals may compare prospective outcomes from current actions. This is evidence that not only *forward* looking expectations (Köszegi & Rabin 2006) may affect reference point formation.

Most of the previous papers on investment decisions with reference–dependent preferences (Berkelaar, Kouwenberg & Post 2004, Gomes 2005, Siegmann 2002, Yogo 2008) do not allow for beliefs to play a role in determining the reference point, and therefore they are usually deterministic with an adjustment based on starting conditions and recent outcomes. Our approach fills this gap.

On the other hand, chapter 4 brings new insights to the literature concerning the allocative efficiency, of two standard auction formats, in the interdependent value auction setting when bidders are asymmetrically informed (Krishna 2003). The experimental findings contribute to the experimental literature of auctions. Most work in this field have focused on either private value auctions or pure common value auctions (Kagel & Levin 2011, 1995).

In what follows I present the chapter on social interactions, heterogeneous beliefs and occupational choice (chapter 2). Then I study reference points formation and investment decisions (chapter 3) and then move to the analysis of auctions in an interdependent value setting with asymmetrically informed bidders (chapter 4). The final chapter 5 provides some concluding remarks and directions of future work

Chapter 2

Occupational choice and social interactions: A Study of Victorian London

"Society is tending more and more to spread into classes, – and not merely classes but localised classes, class–colonies. It is not in London merely, nor as a matter of business, and in consequence of the division of labour that this happens. It is not simply that lawyers dwell with lawyers in the Temple, publishers with publishers in the Row, bankers with bankers in Lombard street, merchants with merchants in Mark lane, carriage-makers with carriage-makers in Long Acre, and weavers with weavers in Spitalfields. But there is a much deeper social principle involved in the present increasing tendency to class-colonies."

— The Economist, June 20, 1857

2.1 Introduction

Despite the wide literature on social interactions and their importance in shaping individual behaviours, in particular labour market decisions (Topa 2011), credibly identifying social interactions remains challenging. First, the researcher must determine the appropriate reference group. Second, unobserved attributes that are correlated between peers, due to self-selection into the group or common information shocks, may generate a problem of confounding variables (*correlated effects*). Third, the reflection

[0]We thank the Archaeology Department of the Museum of London for sharing their geographical referencing of London's historical map. Daniel Felipe Martinez Enriquez assisted us in the extension of this map. The authors acknowledge the use of the UCL Legion High Performance Computing Facility (Legion@UCL), and associated support services, in the completion of this work. We are grateful to all discussants of IFS, UCL and Cambridge-INET Institute and HCEO Chicago Summer School in Social Economics, University of Cambridge for their comments.

problem or the simultaneity in peer behaviour may hinder identification of *contextual effects* (the influence of peer attributes) from *endogenous effects* (the influence of peer outcomes) (Manski 1993). Discrete choice frameworks face further challenges. In particular, the beliefs equilibrium condition cannot be made implicit in the reduced form due to non–linearities. For the same reason, multiplicity of the equilibrium must be taken into account in order to get consistent and efficient estimates.[1]

Our aim in this chapter is twofold: first we provide a discrete choice model addressing these identification issues, and second we offer an empirical application. We present a model of multinomial choice to assess how interaction among members of a group affects their occupational decision. At the core of our specification there is an interplay between an endogenous local network effect and an exogenous effect through public good provision and peer's characteristics. The local network effect represents the externality members of a same group exert on each other. An individual's occupational decision is affected by his rational expectation of his peers' occupational choices while taking into account their characteristics and the strength of their ties. We allow for asymmetric influence while accounting for the network structure introduced by a weighting matrix which captures closeness between each pair. The provision of public goods at a higher level affects the cost of occupation by the characteristics of the area and its residents. We allow for correlated effects at the group level which may reflect information shocks or demand shifts that hit the group as a whole. The nonlinearity introduced by the discrete choice breaks down the linear dependence at the root of the reflection problem, while the asymmetries in the behavioural influence, introduced by proximity weights, allow us to separately identify endogenous and contextual effects from group unobservables. We establish the conditions under which a unique equilibrium can be found. A recursive Pseudo Maximum Likelihood estimation with equilibrium Fixed Point subroutine (Aguirregabiria & Mira 2007) is used in the estimation to solve the problem of indeterminacy due to multiplicity of equilibria in the consistent beliefs condition.[2] In light of our simulations, we conclude that the incidental parameters issue[3] can be dealt with given the nature of our data set, where groups are sufficiently large and there is variation in number of members across them.

We apply our empirical approach to occupational choice in nineteenth century London. For this purpose, we have constructed a new dataset which allows us to geographically locate individuals down to the street level. Matched with the 1881 full census,

[1]See Blume, Brock, Durlauf & Ioannides (2010) for a complete survey of the literature and its challenges

[2]However, we also report the results from the Relaxation Method (Kasahara & Shimotsu 2012) that converges to the true parameters whenever the fixed point constraint does not have local contraction properties in a neighbourhood of the true parameters.

[3]Reminiscent of the non–linear panel data with fixed effects literature (Arellano & Bonhomme 2011).

we are able to determine the occupation of each individual and their characteristics. Victorian London provides a compelling case study. Given that parishes play an important social role in the community and parish membership were based on residency, ecclesiastical parish boundaries dating back to the 17th century provide a convincing proxy for social networks. We rely on the unique two-tier administrative system created by the Metropolis Management Act of 1855 to deal with potential self-selection concerns. Prior to the act, parishes were not only an ecclesiastical and social subdivision, but they were also the districts of local civic government responsible for the administration of taxes in return for many public good services. This Act separated the civil (i.e. dealing with the public good provision) from the social (i.e. fostering social ties) duties of parishes. Smaller parishes were grouped together to form local Board of Works (BW) while larger parishes were elevated to the status of Vestry. BW and Vestries were now in charge of public good provision. In practice, this meant that residents from the same BW living in adjacent ecclesiastical parishes shared the same local institutions but belonged to different social groups. We claim that location decisions were based on BW rather than ecclesiastical parishes. In other words, the Act made group membership orthogonal to other unobservables that affect individual labour market decisions.

From our empirical investigations, we document spatial clustering of occupation in 1881 London. Our results highlight the importance of social networks on occupational choice. Moreover, we uncover how networks have distinct impacts on labour outcomes depending on the type of occupation. Networks have a positive impact for the unemployed and those in industrial occupations while they have a negative impact for those in commercial occupations. There are no social network effects for domestic and professional occupations. Many contextual variables are significant in influencing occupational choice.

It is important to underscore some limitations of our approach. Ecclesiastical parish boundaries might not capture the entirety of a residents social network. Measurement errors and/or misspecification are a concern. Relationships are difficult to observe and quantify. As robustness check, we use pseudo boundaries to test the validity of our social group definition. Given the limitations of our data, we remain agnostics about potential mechanisms driving this spatial patterns. Our results are consistent with models in which agents' employment is affected by information exchanged locally within individuals' social group (Bayer, Ross & Topa 2008, Calvó-Armengol & Jackson 2007). However, other potential channels include social norms or stigma effects (Akerlof 1980), imitation, learning, and complementarities in production.

This chapter contributes to the literature on several fronts. First, we add to the scarce literature on multinomial choice models with social interaction as the only references dealing with a similar framework as far as we know are Brock & Durlauf (2002, 2006) and Bayer & Timmins (2007). Brock & Durlauf (2001) show how contextual and endogenous effects can be identified in a binary choice model with group interactions and no group unobservables. The non–linearity imposed by the logit structure on errors allows them to break the reflection problem documented in the linear–in–means case. Given the information structure of their game, decision makers form rational expectations on other's decision, such belief's structure introduce multiple equilibria. Brock & Durlauf (2006) extend the previous binary decision into the multinomial logit case. They do not allow for group–level unobservables and provide sufficient conditions for identification of the endogenous and contextual effects. Our work is also related to Brock & Durlauf (2007) study on partial identification of binary choice outcomes with group interactions, which relaxes random assignment, known distribution of errors and allows for the presence of group unobservables.

Most studies, including Brock & Durlauf (2002, 2006), follow Manski (1993) to impose rational expectation condition on the subjective choice probabilities of the individual in a large group interaction setting. An individual is equally affected by all the other members in the same group, and he forms rational expectations regarding the choice probabilities of all the other group members. Lee, Li & Lin (2014) incorporates network interactions, as opposed to group interactions, and asymmetric influence in a binary choice model. They allow individual characteristics to enter the information set, so instead of forming rational expectation on the expected behaviour of the group as a whole (i.e. every individual within a group has the same rational expectation) they allow each individual to control for the observed characteristics of other in the group and therefore the rational expectation is a vector of individual choice probabilities of all members in a group. By allowing a network structure and rational expectations in a multinomial choice model we provide a direct extension to Lee et al. (2014) binary choice with asymmetric influence model and Brock & Durlauf (2006) multinomial choice with symmetric influence model.

We also explore the social network structure which few have explored. A notable exception is Lin (2010) who explores various specifications of the spatial weights matrix. To account for the heterogeneity among peers, she allows elements of the weighting matrix to depend on friend nomination order, on the amount of activities associated together, etc. In our setup, we use various measures of geographical distance between the residence of members of the same group. These spatial weights matrix capture the strength and/or availability of contacts.

Second, due to the lack of information on actual contact, researchers usually proxy the relevant group using some arbitrary metric of distance based on social and/or geographical proximity such as school (Gaviria & Raphael 2001, Hoxby 2000), grade (Hanushek et al. 2003), rooms and dorms (Sacerdote 2001), neighbouring families (Bayer et al. 2008, Helmers & Patnam 2014, Solon et al. 2000) and gender or race (Patacchini & Zenou 2012) In the modern world of easy mobility and access to communication technologies, there is a legitimate concern that physical distance nowadays may have become less and less important in shaping social networks. As Manski (2000) has emphasised, presuming we know the true social network is a very strong assumption and may not be plausible in many cases. The absence of a coincidence between measured social groups and true social groups will induce complicated patterns of interdependences in errors across individuals as well as make it difficult to assess counterfactuals such as the effects of changes in the compositions of measured groups. Therefore, focusing on this historical time period provides the advantage that we have a more credible proxy for social networks as interactions were mostly geographic in nature. Moreover, the religious feature of our measure offers an additional relevant dimension to social networks in a period were religion played a central role (Booth 1897).

Third, in the absence of random peer groups (Hoxby 2000, Sacerdote 2001) most studies incorporate group-specific fixed effect and/or group random effects to account for correlated effects. These studies justify this strategy by arguing that, in most contexts, individuals choices cannot narrow their preferences down to the smaller preferred unit. For instance, in the case of class–schools choice, families can somewhat decide which school to send their children but cannot decide which class they should belong to. Thanks to the unique administration layout of the period studied, we provide plausible reasons why adding a fixed effects at the administrative level might better control for self-selection into a group.

Finally, our chapter also contributes to a wider strand of literature interested in evaluating the empirical relevance of the social networks on labour market outcomes. Prior empirical work on the effects of contacts on job finding, and unemployment duration generally confirms that contacts are individually beneficial to workers (Akerlof & Kranton 2000, Blau & Robins 1990, Kramarz & Skans 2007). As Topa & Zenou (2014), we also document spatial clustering of occupations within a city and attempt to distinguish neighbourhood from network effects.

In particular, our chapter hints at an already well-documented observation about the transmission of job opportunities by peers. For instance, Topa (2011) reports that studies commencing in 1970 and using a variety of data sources, find that at least half of all jobs are typically found through informal contacts rather than through formal search methods. Patel & Vella (2013) find that new immigrants are more likely to choose the same occupation previous immigrants from the same country have chosen. Given that social contacts enhance the spread of information, our chapter is also related to the literature on social capital (Knack & Keefer 1997).

The rest of the chapter is organised as follows. Section 2.2 present the multinomial model with social interaction and asymmetric influence. Section 2.3.1 paints the historical background of London in the nineteenth century as the setting of our application while section 2.3.2 presents the the dataset. Section **??** presents our results and 2.4 investigates how robust they are. We finally summarise our findings and conclude in the last section.

2.2 Empirical Model

In this section, we present a model of occupational choice under incomplete information and network interactions. Then we proof that, given our assumptions on locational decisions and information shocks, the structural parameters are identified and finally describe the estimation approach we use. We borrow heavily from Brock & Durlauf (2006) and follow his notation.

2.2.1 Specification of the structural model

We consider a situation where there is a set P of social groups. There is also a set B of administrative areas. For each $b \in B$ there is a collection of social groups P_b belonging to the same area, i.e. $P_b = \{p \in P \mid b(p) = b\}$. Individual $i \in \{1, \cdots, N\}$, characterised by vector $\mathbf{x}_i$ ($dim(\mathbf{x}_i) = K$), belongs to a social group p with n_p members and belonging to administrative area b.[4] Each individual, taking group membership as given, chooses an occupation $y \in \Omega = \{0, 2, \cdots, L-1\}$ earning a market wage ω_y.

In order to capture the potential interaction and/or the strength of ties between individuals, we allow for social interactions to be mediated by the social or spatial distance between each duple $i, j \in p$. For a given reference social group p, we allow for a weighting matrix $\mathbf{W}_p$, with entry $w_{p,ij}, \forall i, j \in p$ measuring the extend to which an individual j influences i's occupational choice, where $w_{p,ii} = 0$. Denote $\mathbf{w}_{p,i}$ as the $1 \times n_p$ row-normalised vector of weights for individual i. Denote individual i's neighbours as $nei_{p,i} = \{j \in p \mid w_{p,ij} \neq 0\}$. Agents therefore may only interact with a subset of individuals identified as his or her peers.

At the core of our specification there is an interplay between an endogenous network effect and an exogenous effect through public goods provision and peer's characteristics.

[4] With a bit abuse of notation we use p both as the social group label and the set of all individuals belonging to that group, i.e. $\mid p \mid = n_p$.

The local effects can be interpreted as a pure network externalities depending on the occupation chosen which is captured by $\phi_y(\omega_y, \mathbf{x}_i, \mathbf{s}^e_{py} \mid \mathbf{W}_p)$. $\mathbf{s}^e_{py}$ is the expectation an individual in group p form on the action taken by any other individual in their group. Therefore, $\mathbf{s}^e_{py} = \left(s^e_{py,1}, \cdots, s^e_{py,n_p}\right)'$ is a vector where entry $s^e_{py,j}$ is the belief any individual in group p has on j taking action y.

$\phi(\cdot)$ may capture the expected future higher/lower benefits of having more peer's on the same occupation due to job information flows or local competition (Anderberg & Andersson 2007, Granovetter 1973) or the availability of handling more efficiently occupation–specific problems when there are more workers of the same type around due to local complementarities or local public goods (Benabou 1993, Kim & Loury 2013).

The cost of following occupation y is given by $C_{b,y}(\mathbf{z}_{p,i})$ where $\mathbf{z}_{p,i}$ is the exogenous characteristics at the group level, $dim(\mathbf{z}_{p,i}) = S$.[5] And refers to the public goods provided by the administrative area or any effects explained mainly by peer's characteristics.

Individual i's decision problem is to choose an occupation y such that

$$\max_y \phi_y\left(\omega_y, \mathbf{x}_i, \mathbf{s}^e_{py} \mid \mathbf{W}_p\right) - C_{b,y}\left(\mathbf{z}_{p_i}\right) + \nu_{p,y,i}, \tag{2.1}$$

where $\nu_{p,y,i} = \epsilon_{y,i} + u_{p,y}$ incorporates preference shocks that depends on individual's decisions $\epsilon_{y,i}$ and a group effect, unobservable to the econometrician, denoted by $u_{p,y}$.

We parametrise $\phi_y(\omega_y, \mathbf{s}^e_{py} \mid \mathbf{W}_p) = k_y + \mathbf{x}_i\mathbf{c}_y + \mathbf{w}_{p,i}\mathbf{s}^e_{py}J_y$. Notice we do not include wages ω_y directly in this approximation. Contrary to common practice, we are not focusing on the effect of foregone earnings influencing occupational choice (Boskin 1974, Heckman & Honore 1990). However, k_y is a constant for each occupation and therefore captures occupation-specific characteristics including wages. The main variable of interest, $\mathbf{w}_{p,i}\mathbf{s}^e_{py}$, is the endogenous social interactions, or the pure network externality. J_y would capture how individual's occupational choice is affected by the belief's on peer's decisions weighted by the strength of ties.

The term $-C_{b,y}\left(\mathbf{z}_{p,i}\right) = \mathbf{z}_{p,i}\mathbf{d}_y + \tau_{b,y}$ represents the opportunity cost of following a certain occupation within an administrative area. It includes the characteristics of the group and administrative area. The former includes characteristics such as the average age, sex, number of children and wealth of residents of the social group. One can think of demand-driven goods or services specific to certain groups which translate in

[5] Among its variables we include $\mathbf{w}_{p,i}\mathbf{X}_p$ where $\mathbf{X}_p$ is the $n_p \times K$ matrix with j-row element $\mathbf{x}_j$.

differences in occupational opportunity cost. One example may be the clustering of the presence of certain firms at specific neighbourhoods. The latter set of variables accounts for public amenities offered at the administrative level which can lead to differences for occupations. Education is an obvious example.

In sum, an individual i who belongs to group p gets utility from choosing y that can be approximated by

$$V(y; \mathbf{x}_i, \mathbf{s}^e_{py}, \mathbf{z}_{p,i}, \tau_{b,y}, u_{p,y}, \epsilon_{y,i}, \mathbf{W}_p) = k_y + \mathbf{x}_i\mathbf{c}_y + \mathbf{z}_{p,i}\mathbf{d}_y + \mathbf{w}_{p,i}\mathbf{s}^e_{py}J_y + \tau_{b,y} + u_{p,y} + \epsilon_{y,i}$$

In Manski (1993)'s terms, J_y is the endogenous effect, $\mathbf{d}_y$ is the contextual effect and $u_{p,y}$ is the correlated effect. The endogenous effect describes how the expected behaviour of peers affect an individual's occupational choice. The contextual effect reflects how the characteristics of fellow group members affects individual i's choice of occupation y. The correlated effect arises through endogenous group formation, common institutional or environmental factors which cause group members to behave similarly even in the absence of social effects.

We assume that

Assumption 2.2.1. $\epsilon_{i,y}$ *are independent and identically distributed across and within groups* p *with known distribution function* F_ϵ*,*

We further assume an individual i does not observe other agents' preference shocks. We therefore have a global interaction model with incomplete information where agents' decisions only depend on their beliefs about other members of the group $\mathbf{s}^e_{py}$.

In Brock & Durlauf (2006) agents within a group possess symmetric influence due to **group interactions**. This is due to every individual being linked to everybody else, attaching equal weight to their influence (i.e. $w_{p,ij} = w \neq 0 \forall j \in p \setminus \{i\}$) and having the same information set ($F_{X_p} \in \mathbb{I}_i$ for all $i \in p$) therefore agreeing on the beliefs about the action of everybody else. However, in the present setup agents only interact with a subset of individuals identified as his or her peers, **network interactions**, and we allow for asymmetric influence mediated by $\mathbf{W}_p$ following Lee et al. (2014).

The intuition of incorporating asymmetric influence is to allow individuals forming beliefs on their peer choices while taking into account their specific characteristics (therefore $\mathbf{x}_j, \mathbf{z}_p$ enters $\mathbb{I}_i$ and not only their empirical distribution, as is the case in Brock & Durlauf (2001)), but also to recognise that some individuals may exert a larger influence on others due to proximity (and therefore, $\mathbf{W}_p$ is also included in $\mathbb{I}_i$).

As standard in the literature (Anderson, De Palma & Thisse 1992, Blume et al. 2010) we assume a Gumbel distribution $F_\epsilon(\epsilon_{y,i} < \epsilon) = \exp(-\exp(\epsilon))$.[6] It follows that agent i, belonging to social group p, chooses occupation y with probability given by

$$\mathbb{P}(y \in \arg\max_{y'\in\Omega} V(y';p,\cdot) \mid \mathbb{I}_i) \equiv s_{py,i} \quad = \frac{\exp\left(k_y + \mathbf{x}_i\mathbf{c}_y + \mathbf{z}_{p,i}\mathbf{d}_y + \mathbf{w}_{p,i}\mathbf{s}^e_{py} J_y + \tau_{b,y} + u_{p,y}\right)}{\sum_{y'\in\Omega} \exp\left(k_y + \mathbf{x}_i\mathbf{c}_{y'} + \mathbf{z}_{p,i}\mathbf{d}_{y'} + \mathbf{w}_{p,i}\mathbf{s}^e_{py'} J_{y'} + \tau_{b,y'} + u_{p,y'}\right)}. \tag{2.2}$$

Under the rational beliefs condition, subjective beliefs on j's occupational choice, $\mathbf{s}^e_{py,j}$, should

a) be agreed upon every individual belonging to the same group

b) such beliefs should match objective beliefs $s^e_{py,j} = s_{py,j}$.

Both conditions imply that the vector $\mathbf{s}_{py}$ is the fixed point solution to the following expression

$$\mathbf{s}_{py} \equiv \begin{pmatrix} s_{py,1} \\ \vdots \\ s_{py,n_p} \end{pmatrix} = \begin{pmatrix} \frac{\exp\left(k_y + \mathbf{x}_1\mathbf{c}_y + \mathbf{z}_{p,i}\mathbf{d}_y + \mathbf{w}_{p,1}\mathbf{s}_{py} J_y + \tau_{b,y} + u_{p,y}\right)}{\sum_{y'\in\Omega} \exp\left(k_{y'} + \mathbf{x}_1\mathbf{c}_{y'} + \mathbf{z}_{p,i}\mathbf{d}_{y'} + \mathbf{w}_{p,1}\mathbf{s}_{py'} J_{y'} + \tau_{b,y'} + u_{p,y'}\right)} \\ \vdots \\ \frac{\exp\left(k_y + \mathbf{x}_{n_p} c_y + \mathbf{z}_{p,i}\mathbf{d}_y + \mathbf{w}_{p,n_p}\mathbf{s}_{py} J_y + \tau_{b,y} + u_{p,y}\right)}{\sum_{y'\in\Omega} \exp\left(k_{y'} + \mathbf{x}_{n_p}\mathbf{c}_{y'} + \mathbf{z}_{p,i}\mathbf{d}_{y'} + \mathbf{w}_{p,n_p}\mathbf{s}_{py'} J_{y'} + \tau_{b,y'} + u_{p,y'}\right)} \end{pmatrix} \tag{2.3}$$

If we collect the $n_p \times L$ matrix $\mathbf{S}_p = (\mathbf{s}_{p0}, \cdots, \mathbf{s}_{pL-1})$ and denote the RHS as $\Psi(\cdot)$ we get

$$\mathbf{S}_p = \Psi(\mathbf{S}_p, \boldsymbol{X}_p, \mathbf{Z}_p, \boldsymbol{W}_p; \boldsymbol{\theta}) \tag{2.4}$$

where $\boldsymbol{\theta} = \left(k_y, c_y, d_y, J_y, (\tau_{b,y})_{b\in B}, (u_{p,y})_{p\in P}\right)_{y\in\Omega}$
This expression could present multiple solutions. The following proposition provides a sufficient condition on J for the existence of a unique equilibrium in the multinomial case with asymmetric influence.

Proposition 2.2.2. ***Multiplicity.*** *In the multinomial choice model with asymmetric influence and network interactions given by 2.1 and 2.4 with $J_y = J$ for all $y \in \Omega$ and abstracting from the effect of the F_X on choice probabilities, if $|J| < 4\left(1 - \frac{1}{L}\right)$ then there is a unique equilibrium.*

[6] Imposing a normalisation on the dispersion in the random utility term equal to 1 and with zero location parameter. This assumption amounts to homoskedasticity. However, due to the spatial nature of our sample we would like to modify this in the future to account for spatial correlation. This would come with an additional requirement: we would have to assume that individuals somehow know the spatial structure of the error term as to form beliefs that are rational.

Proof. Let us leave group p and choice–group unobservables conditioning implicit. Let us assume $\mathbf{Z} = \mathbf{W}\mathbf{X}$. We know that $s_{y,i} = \frac{\exp(k_y + \mathbf{x}_i\mathbf{c}_y + \mathbf{w}_i\mathbf{X}\mathbf{d}_y + \mathbf{w}_i\mathbf{s}_y J)}{\sum\limits_{y'} \exp\left(k_{y'} + \mathbf{x}_i\mathbf{c}_{y'} + \mathbf{w}_i\mathbf{X}\mathbf{d}_{y'} + \mathbf{w}_i\mathbf{s}_{y'} J\right)}$. Therefore

$$m_{y,i} \equiv s_{y,i} - s_{0,i} = \frac{[\exp(g_{y,i} + \mathbf{w}_i\mathbf{m}_y J) - 1]}{1 + \sum\limits_{y' \neq 0} \exp\left(g_{y',i} + \mathbf{w}_i\mathbf{m}_{y'} J\right)} \equiv \psi_{y,i}(\mathbf{M}, \theta, \mathbf{G}),$$

where $g_{y,i} \equiv k_y - k_0 + \mathbf{x}_i(\mathbf{c}_y - \mathbf{c}_0) + \mathbf{w}_i\mathbf{X}(\mathbf{d}_y - \mathbf{d}_0)$, $\mathbf{m}_y = (m_{y,1}, \cdots, m_{y,n})'$ and denote $\boldsymbol{\psi}_y = (\psi_{y,1}, \cdots, \psi_{y,n})'$

We know that the $n \times (L-1)$ matrix $\mathbf{M} \equiv (\mathbf{m}_1, \cdots, \mathbf{m}_{L-1}) = (\boldsymbol{\psi}_1, \cdots, \boldsymbol{\psi}_{L-1}) \equiv \boldsymbol{\Psi}$. If we assume $\mathbf{g}_y = 0 \forall y \in \Omega$ (i.e. we ignore the effect of X on choices) we get that $\mathbf{m}_y = 0 \forall y$ is an equilibrium. To see this, notice that $\boldsymbol{\psi}_{y,i}(0, J, 0) = \frac{\exp(\mathbf{w}_i\mathbf{0}J) - 1}{1 + \sum\limits_{y' \neq 0} \exp(\mathbf{w}_i\mathbf{0}J)} = 0$ for every i, y.

Given this, consider the case for $\hat{\mathbf{M}} = (\mathbf{m}_1, \mathbf{0}, \cdots, \mathbf{0})$ then $\mathbf{m}_1 = \boldsymbol{\psi}_1(\hat{\mathbf{M}}, J, 0) = \frac{\exp(\mathbf{W}\mathbf{m}_1 J) - 1}{L - 1 + \exp(\mathbf{W}\mathbf{m}_1 J)} \equiv \boldsymbol{\psi}_w(\mathbf{m}_1)$. As we want to find a sufficient condition for uniqueness, our aim is to find parameters for which the vector $\boldsymbol{\psi}_w(\mathbf{m}_1)$ is a contraction mapping. Define the metric space $(\boldsymbol{\psi}_w(\mathbf{m}_1), ||\cdot||_\infty)$ where $||A||_\infty$ is the maximum absolute row sum norm of a matrix A given by $\max\limits_i \sum_j |a_{ij}|$. By the contraction mapping theorem we know that if there is a $k \in R$ such that $0 \leq k < 1$ and $||\boldsymbol{\psi}_w(\mathbf{m}_1) - \boldsymbol{\psi}_w(\mathbf{q}_1)||_\infty \leq k||\mathbf{m}_1 - \mathbf{q}_1||_\infty$ then mapping $\boldsymbol{\psi}_w(\mathbf{m}_1)$ has a unique fixed point. By the Mean Value Theorem we also now that for every $\mathbf{m}_1, \mathbf{q}_1 \in [-1, 1]^n$ there is a vector $\mathbf{m}_1'$ that, on an element by element basis, lie in between the former two vectors and such that $\boldsymbol{\psi}_w(\mathbf{m}_1) - \boldsymbol{\psi}_w(\mathbf{q}_1) = \nabla\boldsymbol{\psi}_w(\mathbf{m}_1')(\mathbf{m}_1 - \mathbf{q}_1)$. Applying $||\cdot||_\infty$ on both sides of the previous equality, we get

$$||\boldsymbol{\psi}_w(\mathbf{m}_1) - \boldsymbol{\psi}_w(\mathbf{q}_1)||_\infty \leq ||\nabla\boldsymbol{\psi}_w(\mathbf{m}_1')||_\infty ||(\mathbf{m}_1 - \mathbf{q}_1)||_\infty.$$

Then, we need to find conditions on J such that $0 \leq ||\nabla\boldsymbol{\psi}_w(\mathbf{m}_1')||_\infty < 1$. Notice that $\frac{\partial\boldsymbol{\psi}_w(\mathbf{m}_1')}{\partial m_{1,i}} = 0$ given $w_{ii} = 0$ and $\frac{\partial\psi_{w,i}(\mathbf{m}_1)}{\partial m_{1,j}} = \frac{\exp(\mathbf{w}_i\mathbf{m}_1 J) w_{ij} L J}{L - 1 + \exp(\mathbf{w}_i\mathbf{m}_1 J)}$ for every $j \neq i$.

Therefore, $||\nabla\boldsymbol{\psi}_w(\mathbf{m}_1')||_\infty = \max\limits_i \frac{|J| L \exp(\mathbf{w}_i\mathbf{m}_1 J)}{[L - 1 + \exp(\mathbf{w}_i\mathbf{m}_1 J)]^2} \sum_{j \neq i} |w_{ij}| \leq |J| \frac{L}{4(L-1)}$ Which provides the result.

□

The previous result suggest that the more alternatives individuals face, the less likely multiple equilibrium are. It extends Lee et al. (2014) binary outcome framework into a multiple choice one. With more alternatives the non–linearities in the fixed point condition become less pronounced and multiplicity less pervasive.

We conjecture that allowing for X's affecting choices will enlarge the set of values of J for which an unique equilibrium exists. If choices are affected by observable characteristics, such variables serve as a coordination device thus making multiple equilibria less pervasive. Brock & Durlauf (2007) show this is the case for the binary choice with group–interactions case.

As we detail in section **??** our estimation strategy will need to impose the equilibrium constraint on beliefs in order to get consistent estimates of the structural parameters (see Aguirregabiria & Mira (2007)).

2.2.2 Identification

There are two main threats to identification of the structural parameters $\boldsymbol{\theta}$ of the model 2.1–2.3. First, there is the standard problem of non–random sorting of individuals into the group. Individuals choose which group they would like to belong to (i.e. $p \in P$) and with whom they would like to interact with (i.e. $\mathbf{w}_{p,i}$). The resulting correlation in unobservables among peers can lead to serious bias in the estimation of social interaction among peers in the absence of a research design capable of distinguishing social interactions from these alternative explanations. Second, the presence of correlated effects is a concern created by common unobserved information shocks that hit the group as a whole (Manski 1993). For instance, a group p might face an increase in the demand for certain occupation or have access to better information. If this is not appropriately addressed, one cannot separately identify the exogenous effect from the endogenous effect in the presence of unobserved component under symmetric influence framework (Blume et al. 2010, Brock & Durlauf 2001).[7]

The self–selection problem can be dealt with by operating under random assignment based on observables (Brock & Durlauf 2007, Sacerdote 2001). In our setup, we assume that once we control for the administrative area b's characteristics, individuals are at worst equally inclined to choose any group $p \in P_b$ and, once we control for characteristics at the group p level, $\mathbf{W}_p$ is exogenous.

Assumption 2.2.3. ***Random assignment based on*** τ_b***.***

A.2.1 $dF_{\mathbf{X}|\mathbf{W}_p,\mathbf{z}_p,\tau_b,u_p} = dF_{X|\tau_b}$

[7] The identification is further complicated by the simultaneity problem, also named the reflection problem by (Manski 1993, Moffitt 2001), when studying a linear–in–means. However, the non–linear functional form given by the discrete choice model breaks up the simultaneity problem (Brock & Durlauf 2001) For instance, in the symmetric influence multinomial case without group unobservables, (Blume et al. 2010, Theorem 13) provide sufficient conditions for identification of θ up to a normalization on one of the alternatives.

A.2.2 $\mathbf{W}_p \perp \epsilon \mid z_p, u_p, \tau_b$

We now show that even in the presence of correlated effects, if there is enough variation in the weighting matrix across rows, it is possible to separately identify the endogenous effect from the contextual effect within a asymmetric influence framework. This provides an extension to Brock & Durlauf (2006)'s symmetric influence case. Assume from now on that $\mathbf{z}_{p,i} = \mathbf{w}_{p,i}\mathbf{X}_p$.

Proposition 2.2.4. *Under assumptions A.2.2.1, A.2.2.3, $L > 2$ and the following additional assumptions (AA.)*

AA.1 *Joint support of $(\mathbf{x}_i, \mathbf{w}_{p,i}\mathbf{X}_p)$ is not contained in any linear proper subspace of R^{2K}*

AA.2 *The support of $\mathbf{w}_{p,i}\mathbf{X}_p$ is not contained in any linear proper subspace of R^K*

AA.3 *For each y, there is a group p such that conditional on $\mathbf{W}_p\mathbf{X}_p$, $\mathbf{x}_i$ is not contained in any proper linear subspace of R^K,*

AA.4 *None of the elements of $\mathbf{x}_i$ contains bounded support,*

AA.5 *For each y, across different p groups, $\mathbf{s}_{p,y}$ and $u_{p,y}$ are not constant,*

AA.6 *There is a group p for which $\mathbf{W}_p$ presents sufficiently variation across rows.*

then, for model described by 2.1-2.3, the true set of parameters $\boldsymbol{\theta} \backslash (J_y)_{y\in\Omega}$ are identified up to a normalisation while all $(J_y)_{y\in\Omega}$ are identified.

Proof. Given A.2.2.1, A.2.2.3 and normalising common parameters for $y = 0$ to be 0 (i.e. $k_0 = u_{p,0} = 0, \mathbf{c}_0 = \mathbf{d}_0 = \mathbf{0}$ for every p), we know that, for a given b,

$$\log\left(\frac{s_{py,i}}{s_{p0,i}}\right) = k_y + \mathbf{x}_i\mathbf{c}_y + \mathbf{w}_{p,i}\mathbf{X}_p\mathbf{d}_y + \mathbf{w}_{p,i}(\mathbf{s}_{py}J_y - \mathbf{s}_{p0}J_0) + u_{p,y}.$$

Assume there is another set of observationally equivalent structural parameters $\bar{\boldsymbol{\theta}}$, then it must be the case that

$$\mathbf{x}_i(\mathbf{c}_y - \bar{\mathbf{c}}_y) + \mathbf{w}_{p,i}\mathbf{X}_p(\mathbf{d}_y - \bar{\mathbf{d}}_y) + \mathbf{w}_{p,i}\mathbf{s}_{py}(J_y - \bar{J}_y) + \mathbf{w}_{p,i}\mathbf{s}_{p0}(\bar{J}_0 - J_0) = \bar{k}_y - k_y + \bar{u}_{p,y} - u_{p,y}.$$

Notice that for a given p we have, for every given y, that the right hand side remains constant while there is variation, due to AA.1 and AA.6, on the left hand side. Notice that to avoid perfect collinearity we should impose an exclusion on one social group, call it $p_{(1)}$. Then, for the equality to hold, it must be the case that $k_y + u_{p,y} = \bar{k}_y + \bar{u}_{p,y}$ where we can apply the second part of AA.5 and get $k_y = \bar{k}_y$ and $u_{p,y} = \bar{u}_{p,y}$. Then

$$\mathbf{x}_i(\mathbf{c}_y - \bar{\mathbf{c}}_y) + \mathbf{w}_{p,i}\mathbf{X}_p(\mathbf{d}_y - \bar{\mathbf{d}}_y) + \mathbf{w}_{p,i}\mathbf{s}_{py}(J_y - \bar{J}_y) = \mathbf{w}_{p,i}\mathbf{s}_{p0}(J_0 - \bar{J}_0).$$

Given p, i by AA.5, and considering $L > 2$, for the equality to hold for every y it must be the case that $J_0 = \bar{J}_0$. Which leaves us with

$$\mathbf{x}_i(\mathbf{c}_y - \bar{\mathbf{c}}_y) = \mathbf{w}_{p,i}\left(\mathbf{X}_p(\bar{\mathbf{d}}_y - \mathbf{d}_y) + \mathbf{s}_{py}(\bar{J}_y - J_y)\right).$$

Notice that for AA.2 and AA.3, if we fix a parish p for every y, the previous equality can hold if and only if $\mathbf{c}_y = \bar{\mathbf{c}}_y$. Then it must be the case that

$$\mathbf{w}_{p,i}\mathbf{X}_p(\mathbf{d}_y - \bar{\mathbf{d}}_y) = \mathbf{w}_{p,i}\mathbf{s}_{py}(\bar{J}_y - J_y).$$

We know that AA.4 imply that the LHS is unbounded, but given (2.3) we know that each element $s_{py,i} \in [0, 1]$, so it must be the case that $\mathbf{d}_y = \bar{\mathbf{d}}_y$. As we know by AA.5 that for each y, $\mathbf{s}_{py}$ varies across groups it must also be true that $\bar{J}_y = J_y$. □

The previous results suggests that when collinearity between regressors is ruled out,[8] and one imposes sufficient within variation in at least one parish on choices and characteristics, sufficient between variation across parishes on beliefs and, sufficient within variation in at least one parish on weighting matrix then the structural parameters are identified up to some normalization. In the linear–in–means case, Bramoullé, Djebbari & Fortin (2009) also exploit the weighting (i.e. adjacency) matrix structure, in the shape of intransitive triads between members, for identification. Lee (2007) exploit group size variations to show that separate identification is possible. In such cases, a simultaneity problem must also be dealt with. However, the non–linear functional form of our model breaks up this simultaneity (see Brock & Durlauf 2006) and thus, the requirement on the structure of the network is weaker.

Given that the structural parameters are identified, the next task is to define a consistent estimator for our endogenous effect (Gabrielsen 1978). In the next section we provide the details of such estimator.

2.2.3 Estimation

Denoting $\boldsymbol{X}$ as all exogenous observables specified above and $\boldsymbol{W}$ as the observed spatial weights, the pseudo log-likelihood function, taking $\mathbf{S}$, as the collection of $\mathbf{s}_{py}$ for all p

[8] To avoid perfect collinearity due to the presence of fixed effects at both administrative and social group levels one would have to impose an exclusion restriction in the administrative fixed effects with respect to one administrative level, call it $b_{(1)}$; but also, for every $b \in B \setminus \{b_1\}$ one should add an exclusion restriction with respect to one of its social groups. Otherwise, one could include all social group fixed effects $u_{p,y}$ but one.

and all y, as observed and u as a fixed effect at the group p level, is

$$L_N(\boldsymbol{Y} \mid \mathbf{X}, \boldsymbol{W}, \mathbf{S}; \boldsymbol{\theta}) = \frac{1}{N}\sum_{p\in P}\sum_{i\in p}\log\left[\frac{\sum_{y\in\Omega}\left(\exp\left(k_y + \mathbf{x}_i\mathbf{c}_y + \mathbf{w}_{p,i}\mathbf{X}_p\mathbf{d}_y + \mathbf{w}_{p,i}\mathbf{s}_{py}J_y + \tau_{b,y} + u_{p,y}\right)\mathbb{1}_{[y_i=y]}\right)}{\sum_{y'\in\Omega}\exp\left(k_y + \mathbf{x}_i\mathbf{c}_y + \mathbf{w}_{p,i}\mathbf{X}_p\mathbf{d}_y + \mathbf{w}_{p,i}\mathbf{s}_{py}J_y + \tau_{b,y} + u_{p,y}\right)}\right]. \tag{2.5}$$

The Full Maximum Likelihood Estimator of our discrete choice problem with social interactions and incomplete information is given by

$$\hat{\boldsymbol{\theta}}_{MLE} = \arg\max_{\boldsymbol{\theta}\in\Theta}\left\{\sup_{\mathbf{S}} L_N(\boldsymbol{Y} \mid \mathbf{X}, \boldsymbol{W}, \mathbf{S}; \boldsymbol{\theta}) \text{ s.t. } \mathbf{S} = \boldsymbol{\Psi}(\mathbf{S}, \boldsymbol{X}, \boldsymbol{W}; \theta)\right\}$$

For computational reasons, we follow a recursive Pseudo Maximum Likelihood Estimation procedure with a Fixed Point subroutine (PML/FP) (Aguirregabiria & Mira 2007).[9] For this method to solve the indeterminacy problem due to multiple equilibria we need:

1. within a group, every member agrees on which of all possible equilibria is being played;
2. only stable equilibria emerging in the data;
3. the equilibrium selection to be determined by the data; and
4. a good local contraction properties around the true values.

The first step is to find a consistent estimator for $\mathbf{s}_{py}$, denote it $\left(\hat{\mathbf{s}}^0_{py}\right)_{p\in P, y\in\Omega}$. The second step is to fix $\hat{\mathbf{S}}^0$ and do the PML maximisation using a Newton–Raphson algorithm[10] for any further step $t \geq 1$ such that the estimator at step t is

$$\hat{\boldsymbol{\theta}}^t = \arg\max_{\boldsymbol{\theta}\in\Theta} L_N(\boldsymbol{Y} \mid \mathbf{X}, \mathbf{W}, \hat{\mathbf{S}}^{t-1}; \boldsymbol{\theta}) \tag{2.6}$$

where we replace recursively s^t as the one–step iteration of

$$\hat{\mathbf{S}}^t = \boldsymbol{\Psi}(\hat{\mathbf{S}}^{t-1}, \mathbf{X}, \mathbf{W}; \hat{\boldsymbol{\theta}}^t) \tag{2.7}$$

and keep combining ML iteration with fixed–point updating until $\hat{\theta}^t$ is within a level of tolerance with respect to $\hat{\theta}^{t-1}$.[11]

[9] See Hotz & Miller (1993), Pesendorfer & Schmidt-Dengler (2008) for alternative methods, Pesendorfer & Schmidt-Dengler (2010) for some global conditions in which this iterative procedure fails to converge, with probability approaching 1, to the true parameters, and Kasahara & Shimotsu (2012) for local conditions in which it is known to converge to the true parameter.

[10] This two–step estimator using NR algorithm is efficient (Aguirregabiria & Mira 2007).

[11] The literature provides a couple of alternatives, Bisin, Moro & Topa (2011) suggest implementing

We also implement the Alternative Relaxation Method (i.e. *NPL-Λ algorithm*) where the fixed point iteration is obtained instead by a log–linear combination of $\mathbf{\Psi}(\hat{\mathbf{S}}^{t-1}, \mathbf{X}, \mathbf{W}; \hat{\boldsymbol{\theta}}^t)$ and $\hat{S}^{t-1}$. Specifically, we replace the right hand side of (2.7) by $\Lambda^t = \left\{\mathbf{\Psi}(\hat{\mathbf{S}}^{t-1}, \mathbf{X}, \mathbf{W}; \hat{\boldsymbol{\theta}}^t)\right\}^{\alpha} \hat{S}^{t-1^{1-\alpha}}$ with $\alpha \in \{0.1, 0.8\} \approx 0$. (Kasahara & Shimotsu 2012, see their Prop. 5) argue that even when (2.7) does not have a local contraction property around the true parameters, the Λ^t–mapping does. We find that the results reported here do not change significantly.[12]

As we have allowed for correlated effects at the group level, an additional concern with the PML is that $u_{p,y}$ may induce a incidental parameters problem which lead to the inconsistency of maximum likelihood estimators (Neyman & Scott 1948). This arises because the information about the fixed group effects stops accumulating after a finite number of observations. When groups are very small the incidental parameter problem may become important. In a binary choice network model with small groups the implementation of group fixed effect strategy is not feasible as it introduces too many fixed effect parameters to estimate in the model.[13]

An alternative strategy is to base the estimation on a Conditional ML function that differences out the group fixed effects. Such procedure for the non–linear case was first described by Chamberlain (1980) and resembles the within estimator proposed for the linear-in-inclusive-means case by Lee (2007).[14] One needs to construct a likelihood function that conditions on a sufficient statistic for the incidental parameter.[15] This approach produces a likelihood function that does not depend on the incidental parameters and allows standard asymptotic theory to be applied at the inference stage. The estimator converges to the true parameter as the number of groups increases even if the number of observations per group is small (Andersen 1970). In section 2.3.3 we investigate, by means of montecarlo simulations, how acute such problem may be given the features of our data (i.e. large groups, many groups).

such recursive method for $T = 2$ iterations. Lee et al. (2014) substitute the fixed point updating step with $\hat{S}^t$ being the solution to the fixed point iteration $\hat{S}^t = \Psi(\hat{S}^t, \mathbf{X}, \mathbf{W}; \hat{\theta}^t)$. However, our simulation results, see next section, suggest that the method applied here converges faster to the true parameter.

[12]See section 2.4.

[13]Lee et al. (2014) propose to account for correlated effects by including both the fixed effects at a broader level group and random effect at a group level.

[14]Boucher, Bramoullé, Djebbari & Fortin (2014) is the first empirical application of Lee's results and clarify some of the intuition for identification: That is, individuals with larger outcomes have, by construction, worst peers; positive endogenous effects will therefore decrease the dispersion on outcomes, and will do so at a decreasing rate in group size.

[15]See appendix 2.6.2 for further details.

2.3 Empirical application

To illustrate our approach, we describe the effects of social groups on occupational choice in Victorian London. We first layout the historical background of London in the nineteenth century which justifies our use of ecclesiastical parish boundaries as proxy for social group and our use of particular institutional characteristics for administrative areas. The newly constructed dataset is then presented before turning to the results.

2.3.1 Historical background

An observant pedestrian in London may sometimes see, set into the walls of old buildings at ground level, a small stone with two sets of initials on it, standing for the parishes on either side of the stone. These parish boundary markers were once important to residents as it defined the rights and responsibilities of parishes as a basic administrative unit. Below we attempt to summarise the evolution of the local concept of *parish*, to highlight its civil and social role and how the changes in their boundaries justify our social group definition.

Meaning of parish boundaries

Ancient parishes find their origin in the manorial system. Until the seventeenth century, the manor was the principal unit of local administration and justice. Parish boundaries seem to have been determined by the bounds of the original property (Pendrill 1937). In their beginnings parishes remained largely an ecclesiastical unit. However, in due course, the parish boundaries came to matter a lot to residents as parishes became public good providers. The first significant change in the role of the parish came with the "Poor Law" in 1601. It gave parish officials the legal ability to collect money from rate payers to spend on poor relief for the sick, elderly and infirm - the "deserving" poor. The 1662 Poor Relief Act enabled the creation of "civil parishes", a form of parish which existed solely for specific civil purposes and which had no bearing on ecclesiastical affairs. Parish duties included: to levy a compulsory property-based rate; to put the "undeserving" able-bodied poor to work, whilst punishing those who refused to obey; and to supply outdoor relief to the deserving or impotent poor who were elderly, sick or infirm. During the next century, priests' civil duties, together with an increasing number of other civil duties, were either in the hands of the Justices of the Peace[16] or those of the emerging body of parishioners known as the Vestry.

Due to the fiscal impact of Napoleonic wars and new Corn Laws, during the early 19th century the government was forced to reassess the way it helped the most im-

[16]Justices of the Peace were judicial officers elected or appointed to keep the peace.

poverished members of society. The government's response was to pass a Poor Law Amendment Act in 1834, also referred to the"New Poor Law". The new system was still funded by rate payers, but was now administered by "Unions" – groupings of parishes – presided over by a locally elected Board of Guardians. Each Union had responsible for poor relief by providing a central workhouse for its member parishes. It has to draw upon the economic resources available within its boundaries which included poor rates, parish charities and the creative use of freehold land and commons (Birtles 1999, Webb & Webb 1929). Delimiting boundaries was therefore crucial. In effect, the New Poor Law amalgamated the 15,000 parishes in England and Wales into approximately 600 Poor Law Unions and established a Poor Law Commission in charge of implementing national policies (Besley, Coate & Guinnane 2004).

The Metropolis Management Act of 1855 was a landmark in the history of London's government. Prior to 1855 there was no administrative machinery of any kind responsible for the local government of the metropolis as a whole. All that existed, outside the narrow limits of the City, were about three hundred parochial boards operating under as many separate Acts of Parliament (Firth 1888). The 1855 Act established the Metropolitan Board of Works and empowered it to develop and implement schemes of London-wide significance, perhaps the most well known being the London drainage system and the Thames embankments. The Act also created 15 local Boards of Works (BW) which were groupings of 55 of the smaller parishes together and forced the 23 larger parishes to form Vestries with similar duties as depicted by the figure 2.1.[17] The BW and Vestries were given statutory powers to manage and improve of local facilities such as streets, paving, lighting, drainage and sewerage and elected the members of the Metropolitan Board of Works. Given the redistributive nature of the 1855 Act, the rearrangement of London's local authorities was crucial. This responsibility fell into the hands of Cabinet Minister "who will be able to rearrange the boundaries of London unions at discretion".[18]

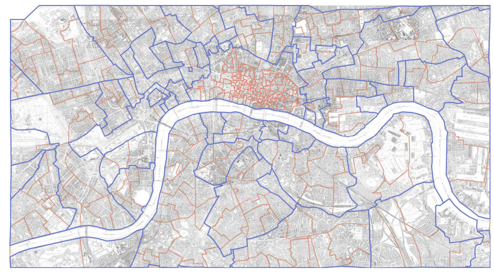
Under the 1855 Act, the boundaries of the ecclesiastical parishes remained unaltered and so was their religious functions. The Compulsory Church Rate Abolition Act of 1868 finally removed the power of ecclesiastical parishes to collect compulsory church rate, from which time they became almost irrelevant as a unit of government. Furthermore, by giving rise to a national system of state education, the Education Act of 1870 (i.e. Forster Act) relieved part of the education role which was previously under the control of the established church. In effect, it created a dual system - voluntary denominational schools and nondenominational state schools. The London School Board,

[17] A detailed list of administrative areas in London can be found in the the appendix **??**.

[18] The Economist, June 19, 1869

covering the whole of London, was created to build and run schools where there were insufficient voluntary school places and to compel attendance.[19]

(a) Ecclesiastical parishes (red)



(b) Board of Works and Vestries (blue)

Figure 2.1: Ecclesiastical and BW borders

Economic and social relevance of parishes

Despite losing importance in terms of civic responsibilities, Victorian age was a religious era and parishes remained an integral part of community life. In his quest to understand the lives of Londoners, Charles Booth dedicated one of its seven volumes to Religious Influences in an attempt to describe the effect of organised religion upon the people of London. In one of his accounts, Booth (1897) stated "so there are other social influences which form part of the very structure of life (...) Among these influences Religion claims the chief part". Such account is corroborated by contemporaneous authors who claim that by the beginning of the nineteenth century "religion was both more pervasive and more central than anything we know in today's Western world" (Friedman 2011).

Anderson (1988) explains how Adam Smith rationalised the economic incentives individuals had to choose to participate in religious activities based on his theory of the capital value of reputation. In particular, he claims that religious membership acted as a club in providing information about individual members' morality which was valuable to reduce transaction costs among them. By providing such reliable information concerning the level of risk attached to dealings with particular individuals, he continues, religious membership improved the efficiency of the allocation of human

[19]The Elementary Education Act 1880 insisted on compulsory attendance from $5 - 10$ years. Elementary education became effectively free with the passing of the 1891 Education Act. The Poor Law remained in force until the 1920s but it gradually lost its functions to other programs and bodies. The administrative division of London was further altered by the Local Government Act of 1888 which created a single London county authority replacing the Metropolitan Board of Works and the Justices of the Peace. These Boards were in turn replaced by the 1903 Metropolitan Boroughs, with similar boundaries to the Boards they replaced. Some workhouses continued in operation until the introduction of the National Assistance Act of 1948. Civil parishes in London were formally abolished in 1965 when Greater London was created, as the legislative framework for Greater London did not make provision for any local government body below a London borough.

resources among their members.

According to Smith (1904), church attendance was not only mandatory but also important to maintain standing within the community.[20]. Church and chapel attendance did not fall between 1851 and 1881, and in absolute terms actually grew up to around 1906, though it fell relative to the population (Smith 1904). In the only reliable Religious Census collected between 1902-1903, 47% of the population in Greater London that could attend a place of worship at least once on a Sunday actually attended.[21]

Membership in a parish was determined by domicile, or by membership in a particular group for which personal parish is established (ethnic parishes, college parishes, etc.). Membership was important as it determined burial, inclusion in the intentions of the Missa pro populo or other spiritual benefits, right to have one's marriage solemnised, etc. Given the fact that religion remained important for residents,"parish boundaries, if they reflected anything, reflected a long-vanished pattern of settlement".(Davis 1988) We therefore base our definition of network on ecclesiastical parish boundaries provided interactions were indeed local in nature.

Interactions restricted by geography

It is usually assumed that for most people in Victorian Britain it was both necessary and convenient to minimise the distance between home and workplace. In the late nineteenth century London the distances over which most people travelled to work remained relatively short. According to Green (1991), this was required both because many trades were casual, and there was thus a strong imperative to be part of a community which knew when work was available (Green 1982, 1991, Hoggart & Green 1991, Johnson & Pooley 1982), and because of the inability of most working people to afford public transport. It was not until after the First World War that the ties between home and workplace were broken, and improved urban transport systems linked to rising real incomes allowed longer-distance commuting for large numbers of people (Dyos 1953, Green 1988, Lawton 1959, Warnes 1972). The mean journey to work for those employed in London was only around five kilometres in the nineteenth century.

[20] "People might attend services on week days if they wished, but it was obligatory on Sundays to join at least in matins and mass, and for at least one member of each family to join in the procession, headed by the priests and clerks with their crosses and banners, that made the perambulation of the church and churchyard. (...) A notorious and unreformed sinner, which would usually mean a heretic who cared nothing for the ways of the Church, would not be allowed to escape by the easy method of staying away. In the tiny parishes religious observance was not only everybody's business, but everybody else's business, and the neighbours would bring him forcibly to the church on Ash Wednesday, where he would be publicly expelled and compelled to come daily to the low side window and listen to mass until Maundy Thursday, when, if repentant, he would be restored" (Smith 1904)

[21] In 1886 a previous census had been carried out providing a larger figure but, given the census was performed in only one day and did not discount for double–counting, it was more imprecise and more dependant on weather conditions of the day.

Professional workers on higher incomes had the longest journeys to work, but in the period 1850 to 1899 professional workers in London still only travelled on average 6.9 km from their home to their workplace. In contrast, skilled manual and craft workers travelled just 3.1 km. Those living within the County of London had especially short journeys to work, for instance residents of East London on average travelled only 2.2 km from their home to their workplace in the period 1850-99. London was notable for the persistence of home working, especially the East End clothing trade. It is estimated that there were over 100,000 home workers in London in 1900 (Schmiechen 1984).

Table 2.1: Distance to work

	Mean journey to work (km)	
Period	Workplace in London	All workplaces
1750 - 1799	2.6	1.7
1800 - 1849	5.1	1.9
1850 - 1899	4.4	2.5
1900 - 1929	10.8	4.3
1930 - 1959	21.0	7.2
1960 +	37.2	14.5
Total sample size	4,957	18,891

Notes: Data extracted from Pooley & Turnbull (1997)

This suggests that social group were probably "local" in nature and a geography–based measure of social group is a plausible assumption let along one that captures a defining social dimension of the time as it was Religion.

2.3.2 Data

We combine several datasets to link the social network and the occupational choice of Londoners. We first use the 100% sample of England and Wales census of 1881 from the North Atlantic Population Project (NAPP). The unit of observation is at the individual level. The census contains the full address of individuals (house number or name, name of street, avenue or road, civil parish and county of residence). In addition to geographic variables, the census also provides a wider range of sociodemographic information: age, gender, place of birth, marital status, number of children, number of servants and family structure as well as information on occupation defined as that in which the individual was principally engaged on the day on which the census was taken (beginning of April). The only economic outcome available in our data is self-reported occupation. There are over 400 occupations such as physician, cook, stable keeper, cabinet maker or farmer.

Using historical maps, we geo reference as precisely as possible all the streets of London. We start from the digitalised map of London dating back to John Rocque's 1746 which was provided by Archaeology Section of the Museum of London. We extend their initial work by manually adding points for each street using the 1882 First Ordnance Survey Map of London. In addition, we locate the church location and record their denomination. We end up with 5998 geographic references to streets or landmarks and 549 churches. Finally, we add the digitalised ecclesiastical parish and and BW/Vestry boundaries provided by the UK Data Service.

In order to geographically locate the individuals in the census on our maps, we use information on place of residence (address, parish and county) from the census and the street points along with the ecclesiastical and BW/Vestry boundaries from the historical map to match these two datasets based on string.

Our final dataset comprises $1,137,876$ individuals for which we can precisely locate down to the street level. This amounts to 70% of matches of the entire population in London in 1881. There are 299 ecclesiastical parishes in Central London and 38 BW/Vestries.

Descriptive statistics

Our sample focuses on native men and women of working age that are household heads (between the ages of 15 and 60). We therefore eliminate foreign-born individuals. We also eliminate individuals who are likely to live in the place where they work such as prisons, workhouses or any other public institution. We finally restrict ourselves to individuals living in parishes for which (i) the BW is composed of at least two ecclesiastical parishes, (ii) with at least 30 residents and (iii) with at least one neighbour living on the same street.[22] We therefore have a total of 200 ecclesiastical parishes within 32 BW. In the appendix 2.6.1, we show the number of ecclesiastical parish per BW, the population density within ecclesiastical parishes, the average number of neighbours per parish and the final areas included in our analysis. Generally, we have large variation across and within BW.

Table 2.2 reports the descriptives statistics of our sample. As expected, men constitute a large fraction of household heads. The mean age is 39 years. The majority of individuals are married with an average of 2 children. The average number of servants, which has been used as a proxy for wealth, is 0.194 with a large variation within the sample. Finally few individuals (13%) have stayed in their parish of birth while 47% have stayed in their county of birth.[23]

[22]These restrictions follow what is standard in the literature and were imposed to avoid noisy estimates whenever there is very few observations.

[23]In Appendix 2.6.1 we map these various characteristics. The south and eastern part of London

Table 2.2: Summary statistics

Average	Unemployed	Professional	Domestic	Commercial	Industrial	Total
male	0.234	0.885	0.407	0.995	0.893	0.835
	(0.424)	(0.319)	(0.49)	(0.074)	(0.309)	(0.371)
age	44.518	39.015	41.680	37.598	38.922	39.243
	(11.239)	(10.408)	(10.733)	(10.335)	(10.548)	10.688
married	0.907	0.890	0.879	0.960	0.944	0.937
	(0.290)	(0.313)	(0.326)	(0.196)	(0.230)	(0.243)
n children	1.810	1.823	1.592	2.028	2.140	2.044
	(1.810)	(1.968)	(1.651)	(1.977)	(2.039)	(1.992)
n servants	0.416	0.605	0.200	0.108	0.158	0.194
	(1.394)	(1.619)	(0.921)	(0.715)	(0.722)	(0.876)
resident p birth	0.094	0.062	0.091	0.116	0.146	0.129
	(0.291)	(0.242)	(0.288)	(0.321)	(0.353)	(0.335)
resident cty birth	0.389	0.337	0.403	0.459	0.494	0.466
	(0.488)	(0.473)	(0.490)	(0.498)	(0.500)	(0.499)
Obs.	10,340	9,559	11,508	28,243	105,464	165,114

Notes: Std. dev in parenthesis. Sample includes only native working-age individuals (between 15 and 60) living in a parish which has a minimum of 30 residents within a BW which has at least two ecclesiastical parishes.

Apart from those unemployed, we have aggregated the remaining occupations into four categories: professional, domestic, commercial, agricultural, and industrial. The employment structure of London was diverse with industrial occupation dominating the labour market. In 1881 6% of the sample were unemployed. Based on our occupational classification, 5.79% worked in a domestic occupation, 6.97% in a professional occupation, 17.11% in a commercial occupation and finally 63% held industrial jobs.[24]

To motivate the choice of nineteenth century London, we map the geographic clustering of occupational choice. In maps of figure 2.2 each panel represent an occupation category. We observe a clear geographical pattern by occupation. Employment appears to be predominant in the central areas of London while unemployment is found the periphery. Professional trades account for a large proportion of West London. Domestic workers are few in East London but more numerous in the City of London and West

are predominantly inhabited by younger, predominantly inhabited by men. Wealthy, captured by the number of servants, is mainly found in the west. The majority of individual living in the south have been born elsewhere.

[24]In the appendix 2.6.1, we provide comparison of these descriptives between the merged and the not merged datasets in order to assess the balance of our sample.

London. In contrast, industrial workers are few in the City of London and West London and more numerous in the East and South. Finally, commercial occupations appear to be more spread out. Agricultural occupation are concentrated in outlying areas of London. However, we see that such concentrations vary both within and across BW (depicted in dotted gray lines) which suggests interactions within social groups may be a driving force behind this striking occupational clustering of labour outcomes across London areas.

2.3.3 Simulation to study the performance of the estimator

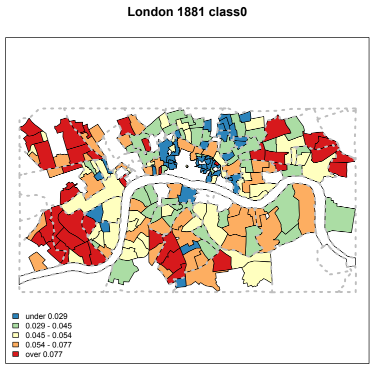
We take the empirical distribution of parish members and minimum number of neighbours living on the same street from our data set (see figure 2.11 in the Appendix) and randomly draw a duple $(n_p, \min_{i \in p} nei_i)$. We then simulate geographic points on $[0, U_{[1,2]}]^2$. We define $w_{p,i}$ as all those individuals that are within the radius δ_p close to i where δ_p is the minimum distance such that every $i \in p$ has at least one neighbour. Individuals face the utility function given by equation 2.1,where we assume $K = S = 1$ and have to choose among five alternatives ($L = 5$).[25]

We assume the true coefficients are given by $J^0 = (3.3, 2.5, 2, 3.2, 3.6)'$ (the endogenous effect), $k^0 = (-1.5, -1.3, -2.4, \ -1.7, -2.1)'$ (alternative–specific characteristics), $c^0 = (1.5, 1.4, 2.1, 0.9, 1.1)'$ (individual characteristics) and $d^0 = (2.2, 2.4, 2.7, 2.3, 2.6)'$ (the contextual effect). Regarding the correlated effect, we assume $u_{p,y} \sim \sigma_y N(0, 1)$ where $\sigma^0 = (0.13, 0.08, 0.18, 0.05, 0.1)$.[26] To simulate optimal choices while imposing consistent asymmetric influence we solve, iteratively, for the belief fixed point of 2.3 until no individual's beliefs change. This guarantees only stable equilibria to emerge in the observational data.

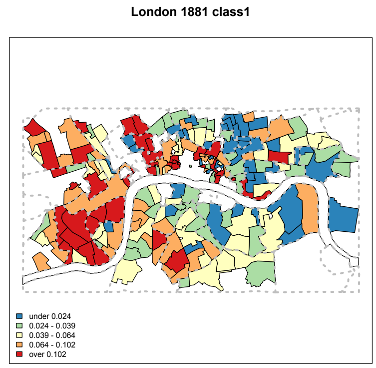
We then assume that the Data Generation Process (DGP) is identical across different groups that belong to world c. We generate 100 "worlds" with $\mid P \mid= \{5, 10, 50, 100\}$ groups each. Even though the DGP is exactly the same for a given world c, w_p^c and X_p^c are random so we have different choices across groups. In the left hand side panels $(a,\ c)$ of Figure 2.3, we depict two examples of different individuals' locations within a group. Red lines represent the links within them (i.e. w_p) and the shape of the point represents the alternative being chosen by each individual (i.e. y_p). On the right hand side panels $(b,\ d)$ we plot the fixed point convergence of hetereogenous beliefs (i.e.

[25] We should make clear that for the simulation these are abstract alternatives and are not related to the occupational choice problem we are interested in. The purpose of the simulation is therefore only to get an idea on how close to true parameters our estimates, following the proposed strategy, could be.

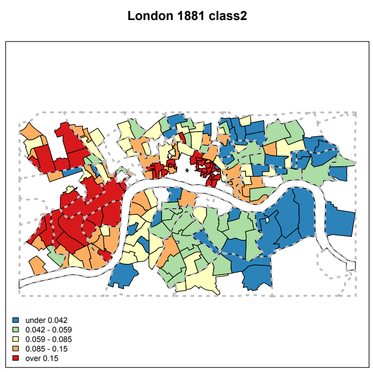
[26] The coefficients were chosen so as to have some alternatives being chosen by less than 10% of the population. In the appendix we include some variations of these coefficients. The results remain valid (see Appendix tables 2.12 and 2.13)



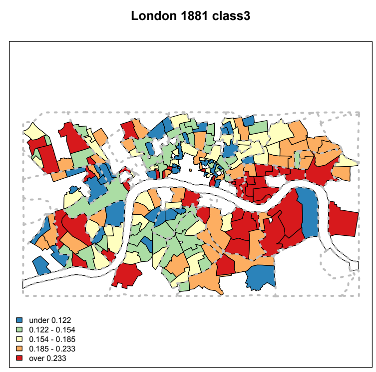
(a) Unempoyed



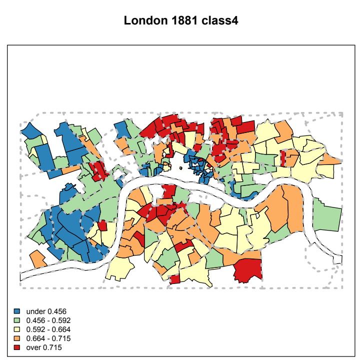
(b) Domestic Occupation



(c) Professional Occupation



(d) Commercial Occupation



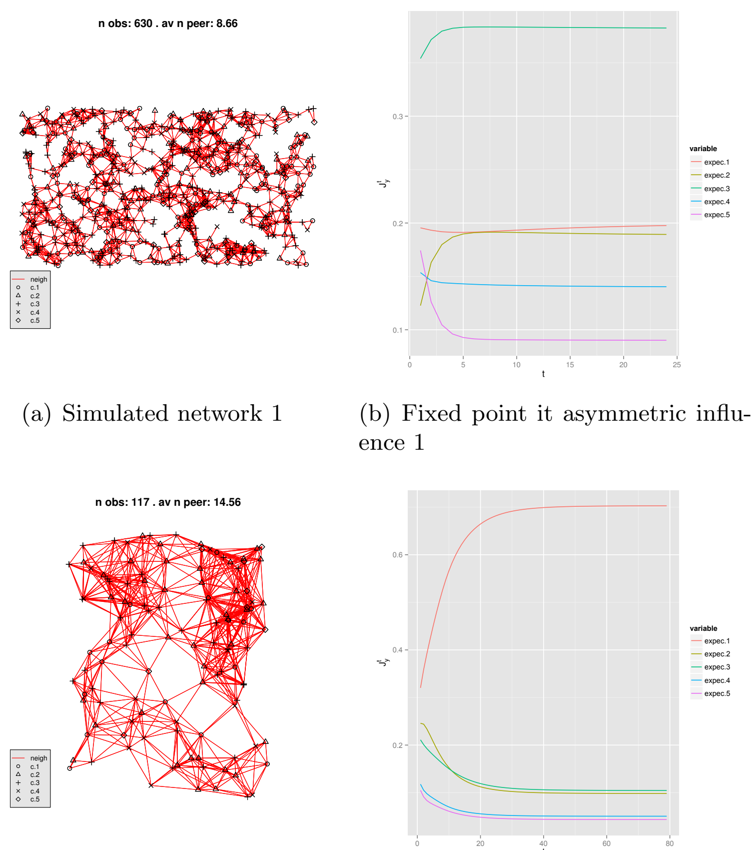
(e) Industrial Occupation

Figure 2.2: Occupations per ecclesiastical parish.

$s_{p,y}$) for our simulated data associated to each group. It is noticeable that differences in the simulated X's may imply variation in $s_{p,y}$ across groups which is necessary for the identification of the endogenous effect (see proposition 2.2.4).

The simulations provide us with observational data on choices $\mathbf{Y}$, characteristics $\mathbf{X}$ and network $\mathbf{W}$ for every group $p \in c$. We perform the estimation by the PML/FP described in equations 2.6-2.7 using a Newton–Raphson algorithm.

This exercise allows us also to investigate how important is the incidental parameter problem and is closer in nature to our real set–up, where we have a city (i.e. London 1881) with 299 different social parishes. Among which 277 of them have $n_p \geq 30$ and 200 for which inhabitants are fully geographically located on our map and belong to Unions composed by at least 2 different social parishes.



(a) Simulated network 1

(b) Fixed point it asymmetric influence 1

(c) Simulated network 2

(d) Fixed point it asymmetric influence 2

Figure 2.3: Some simulated network and beliefs fixed iteration per parishes with θ^0

In Table 2.3 we show the endogenous parameter J_y estimates from the ML estimation assuming beliefs $s_{p,y}$ are not observed. We also explore the effects of measurement errors in the networks. In *Panel A*, the true network w_p is observed, while in *Panel B* the network is only partially observed w_p^* (we assume a truncated version of the true network where up to 10 peers are observed).[27] The last row shows the percentage of individuals choosing each alternative.

Focusing on *Panel A*, we see that the larger the number of parishes is the closer the estimates to the true parameters and the smaller the dispersion. One important result is that estimates of the endogenous effect for a given alternative is very precise when their is a large mass of individuals choosing that alternative. In our case, the estimate for J_4 is imprecise due mainly to very few observations choosing such alternative (i.e. less than 8%).

The results in *Panel B* are not as encouraging which should come as no surprise: as documented in the literature, whenever the true network is not observed, estimates are biased. A truncated network (in this case observing only up to observing 10 neighbours) generally lead to the underestimation of the endogenous parameters. However, a large number of parishes alleviates the underestimation. Consequently, our recursive PML/FP method cannot recover the true beliefs (see figure 2.14 in the appendix). This reinforces the importance of appropriately measuring the social group.

[27]In the Appendix 2.6.3 we include corresponding estimates for the effect of an individual's own characteristics c (Table 2.11) and the contextual effect d (Table 2.10) effects as well as the estimates when both $s_{p,y}$ and w_p are observed (Table 2.9).

Table 2.3: Montecarlo simulations and estimation of J when $s_{p,y}$ is not observed

	(1)			(2)			(3)			(4)			(5)				
	$J0 = 3.3$			$J1 = 2.5$			$J2 = 2$			$J3 = 3.2$			$J4 = 3.6$			$\min nei_i$	n_p
$\mid P \mid$	mean	med	sd	mean	med	sd	mean	med	sd	mean	med	sd	mean	med	sd	mean	
A. PML/FP, w observed																	
5	2.64	3.20	1.82	2.06	2.32	1.36	1.92	1.99	0.99	1.05	2.81	5.46	2.06	4.80	6.75	8.2	4471.0
mse		*3.71*			*2.02*			*0.98*			*34.14*			*47.45*			
10	3.25	3.30	0.45	2.34	2.45	0.53	1.88	1.97	0.59	2.47	3.18	2.65	1.82	3.64	5.17	7.3	8478.6
mse		*0.20*			*0.30*			*0.36*			*7.50*			*29.63*			
50*	3.28	3.29	0.08	2.45	2.48	0.20	1.99	2.01	0.21	3.14	3.16	0.55	3.32	3.73	1.93	7.3	42994.5
mse		*0.01*			*0.04*			*0.04*			*0.31*			*3.76*			
100**	3.30	3.29	0.05	2.44	2.45	0.13	1.99	1.98	0.15	3.15	3.18	0.40	3.29	3.35	1.39	7.8	85752.1
mse		*0.00*			*0.02*			*0.02*			*0.16*			*2.00*			
B. PML/FP, w truncated																	
5	2.33	3.08	2.24	1.61	1.90	1.38	0.40	0.33	0.84	0.65	1.27	3.12	-0.11	0.92	5.45	8.2	4471.0
mse		*5.90*			*2.68*			*3.24*			*16.10*			*43.16*			
10	3.23	3.30	0.50	1.96	2.16	0.70	0.45	0.47	0.55	1.05	1.00	1.82	0.86	1.33	3.73	7.3	8478.6
mse		*0.25*			*0.77*			*2.71*			*7.90*			*21.25*			
50*	3.33	3.34	0.15	2.14	2.17	0.31	0.44	0.43	0.22	1.40	1.50	0.75	1.54	1.72	1.54	7.3	42994.5
mse		*0.02*			*0.23*			*2.47*			*3.79*			*6.62*			

Continued on next page

	$J_0 = 3.3$			$J_1 = 2.5$			$J_2 = 2$			$J_3 = 3.2$			$J_4 = 3.6$			$\min nei_i$	n_p
100**	3.37	3.37	0.10	2.18	2.21	0.21	0.44	0.46	0.17	1.27	1.34	0.51	1.52	1.62	1.23	7.8	85752.1
mse		*0.02*			*0.15*			*2.46*			*4.00*			*5.82*			
choices		*0.27*			*0.30*			*0.27*			*0.09*			*0.08*			

* 105 simulations, ** 107 simulations. *med*: median, *sd*: standard deviation, *mse*: mean square error.

In the Figure 2.4 we depict the correlation and density between the true beliefs spatially weighted (i.e. $w_p s_{p,y}$ denoted as *expw* in the figure) and the estimated beliefs through the PML/FP (i.e. $w_p \hat{s}_{p,y}$) denoted as *sw* in the figure). It is evident that, after a fairly low number of iterations ($T = 58$), the estimated beliefs converge to the true beliefs, which suggests that the estimation procedure leads to a good fit with the real unobserved beliefs.

Taken together, our results suggest that the estimation is generally accurate whenever the true network is observed. Even though we have reasons to belief that in the nineteenth century geographic measures of reference groups where meaningful and religion was an important dimension of social identity, in the robustness section we investigate how sensible our results are to different definitions of social group.

The results suggest that, given the nature of our sample (i.e. generally large reference groups with large number of neighbours), the incidental parameter problem can be downplayed and estimates following the method described in equations 2.6-2.7 converge to the true parameters.[28]

2.3.4 Group interactions and symmetric influence

We now move on to our true data set. As a benchmark we present the estimation for the standard symmetric influence case with group interactions and no correlated effects (Brock & Durlauf 2001, 2006). In the next section, we relax these restrictions allowing for both asymmetric influence due to network interactions and unobservables at the group level affecting individuals' decisions.

The symmetric influence assumption implies that the number of agents in each parish p is sufficiently large so that each agent dismisses his own effect on others' decisions. The condition for rational expectations (2.3) is now given by

$$s_{p,y} = \int \frac{\exp(k_y + x_i c_y + \bar{x}_p d_y + J_y s_{p,y} + \tau_{b,y} + u_{p,y})}{\sum_{y' \in \Omega} \exp\left(k_{y'} + x_i c_{y'} + \bar{x}_p d_{y'} + J_{y'} s_{p,y'} + \tau_{b,y}\right)} d\hat{F}_{\boldsymbol{x}|p}, \text{ for all } p \in P \tag{2.3a}$$

where individuals know $\hat{F}_{\boldsymbol{x}|p}$, the empirical within-group distribution of $(x_i, \bar{x}_p)$. Notice that this expression is no longer a vector value function, but still can present multiple consistent beliefs (Blume et al. 2010).

[28] Even though the consistency properties of such estimator are not analysed here, the result is reminiscent of the importance of rich number of groups and variation in their sizes for consistency of the method proposed by Lee (2007) in the linear–in–means case, which is clarified further in Boucher et al. (2014).

We begin our econometric findings with the estimation of equation (2.5) for the symmetric influence case where $\mathbf{W}_p$ is a matrix with zero along the diagonal and $1/(n_p - 1)$ off–the–diagonal whenever two individuals live in the same parish. We also consider s_p^0 as being equal to the observed weighted average decision at group p ($\widehat{s}_{y,p}^0 = \frac{1}{n_{p-1}} \sum_{j \in p} \mathbb{1}[y_j = y]$). We shy away from allowing any correlated effects, therefore $u_{p,y} = 0 \forall p \in P, y \in \Omega$. We call this first exercise the naive estimates for θ such that

$$\theta^{naive} = \arg \max_{\theta} L(\boldsymbol{Y} \mid \mathbf{X}, \widehat{s}; \theta) \tag{2.8}$$

As previously, individuals choose among five different occupations: "out of the labour force" ($y = 0$), "professional" ($y = 1$), "domestic" ($y = 2$), "commercial" ($y = 3$), "industrial" ($y = 4$). We will use $y = 0$ as the benchmark. Individual characteristics x_i include age, sex, marital status, number of children, number of servants and resident in parish of birth. Group level (i.e. ecclesiastical parish) characteristics $\bar{x}_p = \frac{1}{n_{p-1}} \sum_{j \in p} x_j$. Blume et al. (2010) discuss under which assumptions this estimator is consistent.

Tables 2.4 presents the endogenous effects J_y. Tables 2.14 in the appendix present the other estimates. The first three columns illustrate the naive estimates for J_y using equation (2.8) while the last column uses the PML/FP using equation (2.3a). We successively include sets of variables: the first column includes only individual characteristics x_i, the second column adds to this specification the contextual variables at the social group level $\bar{x}_p$, and the third and fourth columns add the fixed effects at the BW. This is our preferred specification due to its dealing with the self selection problem.[29]

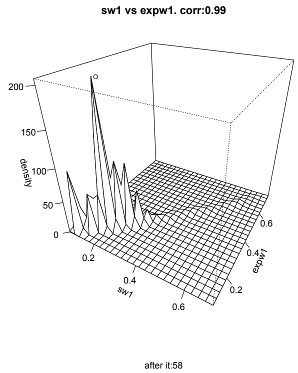
It is immediately apparent from table 2.4 that the contextual effect is an important source of upward bias in our endogenous effect. Groups may differ in average level of schooling, cognitive functioning, occupational structure and wealth level. Moreover, including the BW–occupation–specific dummies shows that there are local factors which play a role on endogenous effects. Comparing the last two columns reveals that the PML/FP procedure leads to smaller parameters which may indicate more accurate estimates due to the fact the specification is now internally consistent with beliefs and the included variables (i.e. weighted estimated beliefs, $\frac{1}{n_{p-1}} \sum_{j \in p} s_{y,p}^t$) are smoother than the non–parametric consistent estimator s_p^0. For this very same reason the log-likelihood values are not necessarily compara

The coefficients for each occupation category are consistent with the presence of local peer effects. The presence of peers in a given occupation has a significant and positive effect on the likelihood of following that same occupation. This is true for all

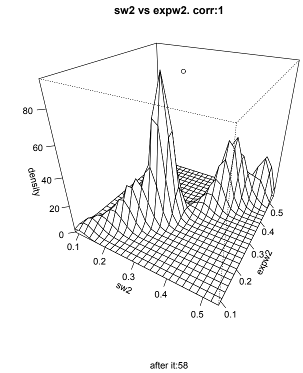
[29] In Figure 2.15 panel *a)* we show the convergence of endogenous coefficients per iterations.

occupations except for commercial ones. Individuals out of the labour force present the largest endogenous effects. The larger the number of unemployed one interacts with, the less likely s/he is to be employed. The present of significant endogenous effects among professional occupations is more puzzling as one expected high–skilled networks to be less geographically restricted. Commercial occupations on the other hand do exhibit significant endogenous effects.

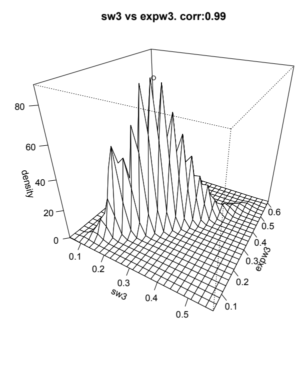
Individual characteristics have the expected sign (see table 2.14). Age affects negatively the propensity of being in any given occupation compared to being out of the labour force. Given that migration might be due to job prospects, we find that those who have not moved away from their parish of birth are more likely to be in a productive occupation. A large number of children decreases the chances of individuals being in an occupation, which could indicate poverty traps or child labour.



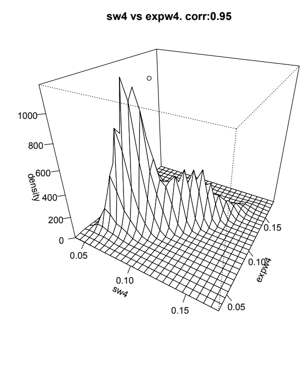
(a) $cor(w_p s_{p,0}, w_p \hat{s}_{p,0})$



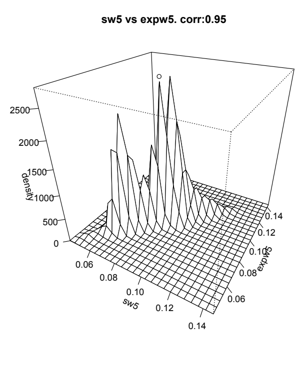
(b) $cor(w_p s_{p,1}, w_p \hat{s}_{p,1})$



(c) $cor(w_p s_{p,2}, w_p \hat{s}_{p,2})$



(d) $cor(w_p s_{p,3}, w_p \hat{s}_{p,3})$



(e) $cor(w_p s_{p,4}, w_p \hat{s}_{p,4})$

Figure 2.4: Density and correlation between true beliefs ($w_p s_{p,y}$) and PML/FP estimates beliefs ($w_p \hat{s}_{p,y}$) after 58 iterations across 5 alternatives

Table 2.4: Estimation of endogenous effects J_y with symmetric influence

vars	**Naive estimation**			**PML/FP**
	(1)	(2)	(3)	(4)
unemployed	10.55^{***}	14.613^{***}	10.044^{***}	5.676^{***}
	(0.363)	(0.55)	(0.791)	(1.643)
domestic	9.522^{***}	10.118^{***}	8.247^{***}	4.059^{***}
	(0.279)	(0.357)	(0.418)	(0.955)
professional	8.151^{***}	6.39^{***}	5.682^{***}	3.773^{***}
	(0.173)	(0.255)	(0.287)	(0.511)
commercial	4.644^{***}	4.63^{***}	2.849^{***}	-0.447
	(0.134)	(0.167)	(0.282)	(2.023)
industrial	1.995^{***}	1.987^{***}	2.414^{***}	2.725^{***}
	(0.071)	(0.1)	(0.142)	(0.386)
log-like	-156320	-156070	-155760	-156820
obs		165114		
x_i	yes	yes	yes	yes
$\bar{x}_p$	no	yes	yes	yes
$\tau_{b,y}$	no	no	yes	yes
$u_{p,y}$	no	no	no	no

Notes: Standard errors in parentheses. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.
unemployed $s_{p,0}$; professional $s_{p,1}$; domestic $s_{p,2}$; commercial $s_{p,3}$; industrial $s_{p,4}$.

2.3.5 Network interactions and asymmetric influence

In table 2.5 we include the asymmetric influence estimation. In this case we define $\mathbf{W}_p$ as matrix with zero diagonal and entry $w_{p,ij} = 1/|nei_{p,i}|$ if j is a *neighbour* of i. We define *neighbour* as any two individuals living within a 50 *mts* radius from each other.[30] We focus on the specifications that controls already for individual characteristics x_i and network–level covariates $\mathbf{w}_{p,i}X_p$. We present the naive estimation (i.e. without imposing equilibrium condition (2.4)) in the first two columns. We perform the PML/FP structural estimation described in expressions (2.6) and (2.7) in the last

[30] As already explained above, we are able to locate individuals down to the street level, therefore the distance between two individuals is taken from the mass point of the streets in which they live. The results that follow hold for different definitions of neighbours, see section 2.4.

two columns. We successively add parish fixed effect to the basic specification.

We can now gain insight into the correlated effects at the social group (i.e. ecclesiastical parish) level. Such correlated effects can take the form of local industries or an inspiring priest which might encourage his parishioners to work or share information. Failing to take the correlated effects into account can lead to serious upward bias.

Comparing column (4) of table 2.4 with column (3) of table 2.5 reveals the difference between symmetric and asymmetric influence. We note that the symmetric influence specification overestimates endogenous effects for out of the labour force, domestic and commercial occupations; while it underestimates the social effect on professional and industrial occupations. Under the asymmetric influence specification, we see that commercial occupations are not subject to positive endogenous effects but instead, negatives ones. Again, the magnitude of the endogenous effect appears very high for professional occupations. In the symmetric case such estimate may be capturing neighbourhood effects rather than social interactions.

Our preferred specification is depicted in the last column. Networks play a significant and positive role for individuals out of the labour force and in industrial occupations. If you expect your peers to be unemployed you are less likely to receive information about job opportunities through informal channels, therefore you are more likely to be unemployed as well (Calvó-Armengol 2004). Computing the marginal effect[31] we find that a one standard deviation change in the weighted expected ratio of unemployed peers leads to a 0.54% increase in the likelihood of being unemployed. Notice that such magnitude is somewhat lower than contemporary studies. Topa (2001) who finds that a one standard deviation of peers' employment leads to a increase in the likelihood of being employed that lies between $[0.6\% - 1.3\%]$, while Bayer et al. (2008) estimates lie somewhere between $[0.8\% - 3.6\%]$.

Industrial occupations are mainly demand–side driven and information about job opening should therefore be easily transmitted. The marginal effects suggest that a one standard deviation change in the peers industrial expected ratio leads to an increase on 8.04% in the likelihood of being employed in a similar occupation. Commercial occupations on the other hand might be more competitive and individuals may want to keep private information on customers or alike for themselves. A one standard deviation increase in peers commercial expected occupational choice reduces the chance of following a similar occupation in 2.46%.

At both end of skill's distribution (i.e. domestic and professionals) we find that our network measure do not explain occupational choice once we allow for unobservables

[31] As pointed out by Lee et al. (2014) one should account for the effect on equilibrium conditions of a covariate change. To compute marginal effects we use the formulas found in appendix 2.6.5

hitting the group as a whole. The forces driving these result might be very different. In the domestic case, the availability of such posts may be very locally restricted. Individuals will tend to live where they work and not necessarily where their perceived peers will interact and share information. Local interactions may not be the channel through which one could hear about such job offers. On the other hand, professional occupations may have two unique features: Firstly, they may be particular prone to locate in particular parishes (probably wealthy ones) compared to other occupations. Thus, once we allow for group unobservables at the parish level (which may account for such self-selection at the parish level), the seemingly large endogenous effect becomes insignificant. Secondly, the professional class is arguably the less spatially confined, and therefore, their social networks may extend beyond a geographical/religious dimension.

The results from the naive and PML/FP estimation are substantially different so it is worth understanding why the latter may be more reliable. We know the naive estimation consistency depends largely on how accurate the local average of occupations incidence is as a proxy for rational beliefs. On the other hand, even with a poor starting estimate on the beliefs, the recursive PML approach may get, after suitable iterations, consistent estimates for s_p (Aguirregabiria & Mira 2007).

Table 2.5: Estimation of endogenous effects J_y with asymmetric influence

vars	**Naive estimation**		**PML/FP**	
	(1)	(2)	(3)	(4)
unemployed	4.667^{***}	3.233^{***}	2.164^{**}	3.057^{***}
	(0.288)	(0.267)	(0.751)	(0.760)
professional	4.857^{***}	2.373^{***}	-0.663	-0.976
	(0.169)	(0.166)	(0.740)	(0.650)
domestic	4.614^{***}	3.991^{***}	3.841^{***}	-0.669
	(0.131)	(0.140)	(0.458)	(1.220)
commercial	1.865^{***}	1.764^{***}	-3.992^{***}	-3.429^{**}
	(0.108)	(0.101)	(1.309)	(1.143)
industrial	2.419^{***}	3.426^{***}	3.342^{***}	3.639^{***}
	(0.067)	(0.071)	(0.252)	(0.265)
log-like	-152620	-152010	-156240	-154100
obs	165114			
x_i	yes	yes	yes	yes

Continued on next page

	(1)	(2)	(3)	(4)
$w_p X_p$	yes	yes	yes	yes
$\tau_{b,y}$	yes	no	yes	no
$u_{p,y}$	no	yes	no	yes

Notes: Standard errors in parentheses. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. Beliefs unemployed $s_{p,0}$; professional $s_{p,1}$; domestic $s_{p,2}$; commercial $s_{p,3}$; industrial $s_{p,4}$.

We also investigate whether there is multiple equilibria given our parameters. Per parish p we then want to find the roots to the large $(n_p(L-1))$–system of non–linear equations described by

$$F(s_p; \hat{\theta}) \equiv \text{vec}(s_p) - \text{vec}(\Psi(s_p, \boldsymbol{X}, \boldsymbol{W}; \hat{\theta})) = 0$$

To do so we follow spectral methods[32] to solve for possible multiple roots. However, for none of the 200 parishes we could find (after using 1005 different starting values per each parish) more than one equilibrium. Even though some of the absolute value estimated parameters in column (4) are above the threshold 3.2 suggested by proposition 2.2.2, we know that in the proposition we shy away from allowing $(k, c, d) > 0$. As pointed out by Brock & Durlauf (2006), for the symmetric case, the presence of individual differences across covariates may increase the threshold for which unique equilibrium exists.

2.4 Robustness checks

We perform several additional results to study how robust our estimates are and also provide evidence that our identifying assumption are likely to hold.

2.4.1 Change of $\mathbf{W}_p$ and placebo coordinates

In the first three columns of table 2.6 we modify the definition of our weights. In column (1) we define our weighting matrix as $w^{\delta}_{p,ij} = 1/|nei^{\delta}_i| \forall j \in p$ such that $|| \, i - j \, || \leq \delta mts$, where index i is used as a label for an individual as well as his coordinates. We use $\delta = \{0, 100\}$. Similarly we follow the same estimation procedure but truncating the number of neighbours to 10, $w_p^{|nei|<10}$.

What we learn from such exercises is how sensitive results are to different definitions of

[32] See Varadhan & Gilbert (2009) for an implementation of such algorithm in *R–package*.

the network. For a distance of 0 *mts* results are very similar to our original same–street interactions. Once we allow for larger radius within the same parish (i.e. 100 *mts*) we see that the endogenous effect becomes significant for professionals. On the other hand, a 10–peers truncated network imply generally underestimated effects.

We also implemented a placebo test in which we randomly allocate individuals on to different streets across all city. One concern is that the aggregation method we are pursuing could, somehow, influence the statistical significance of the results. Such placebo test could shed some light on how important this concern is. Then we followed exactly the same PML/FP estimation as before. The estimates from such placebo test are included in column (4). What we observe is that the endogenous effect is now insignificant for all occupations. It implies that the endogenous effect is not driven by the type of aggregation we used. To also rule out that the endogenous effect found in the previous section is driven by other unobserved geographic characteristics we randomly allocate individuals on different streets *within* the same parish. Results are presented in column (5), it reassures that our estimates of the endogenous effect are mainly driven by network–interactions as opposed to other neighbourhood unobservables.

In columns (6) and (7) we modify the sample. In column (6) we restrict the sample to only individuals that are living within the same county where they were born, in column (7) we restrict our analysis to the younger cohort (i.e. ages within 15 and 30 years). We notice that results change significantly. For non–movers, there are now no significant effect of endogenous effects on unemployment, however the ones on commercial and industrial occupations remain. This suggest that there is migration responding to lack of job opportunities. The case of the younger cohort indicates that the initial occupation of such population tend to be mainly on industrial and domestic tasks. Taken together, these two results suggest an additional heterogeneity that our empirical model is not addressing. Further research is needed to be able to incorporate multiple types and thus, heterogeneous beliefs, into the estimation procedure while accounting for consistent beliefs.

Table 2.6: Robustness checks estimation of endogenous effects J_y with asymmetric influence

vars	**PML/FP**						
	$w_p^{\delta=0}$	$w_p^{\delta=100}$	$w_p^{\lvert nei\rvert<10}$	Placebo coords		Non	Ages
				all city	within parish	movers	15-30
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
unemployed	3.10***	3.577***	2.154***	-2.420	0.476	1.228	1.724

Continued on next page

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
	(0.73)	(0.854)	(0.608)	(19.49)	(2.058)	(1.13)	(1.795)
professional	-0.796	0.077	0.135	-3.927	2.144	-1.429	-0.255
	(0.624)	(0.718)	(0.563)	(12.957)	(2.004)	(0.938)	(1.251)
domestic	-1.545	2.476**	-2.943**	-12.754	0.306	-0.925	2.449**
	(1.075)	(0.936)	(0.953)	(11.74)	(1.425)	(1.55)	(0.773)
commercial	-3.356**	-3.129*	-2.229**	4.436	-2.766	-3.323*	-0.807
	(1.061)	(1.472)	(0.791)	(3.62)	(2.259)	(1.346)	(1.486)
industrial	3.519***	3.351***	2.903***	-4.198	0.732	3.457***	2.843***
	(0.257)	(0.301)	(0.243)	(3.623)	(1.059)	(0.396)	(0.435)
log-like	-154060	-154070	-154340	-183220	-155000	-64636	-37864
obs		165114			165114	76643	42497

Notes: Standard errors in parentheses. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. unemployed $s_{p,0}$; professional $s_{p,1}$; domestic $s_{p,2}$; commercial $s_{p,3}$; industrial $s_{p,4}$. $(x_i, w_p X_p, u_{p,y})$ always included

2.4.2 Alternative relaxation method

As stated above, we also implemented the Kasahara & Shimotsu (2012) NPL–Λ algorithm that converges to the true parameters whenever the fixed point constraint (2.7) does not have local contraction properties in a neighbourhood of the true parameters. Specifically, we replace the right hand side of the fixed point iteration by expression

$$\Lambda^t = \left\{\boldsymbol{\Psi}(\hat{\mathbf{S}}^{t-1}, \mathbf{X}, \mathbf{W}; \hat{\boldsymbol{\theta}}^t)\right\}^{\alpha} \hat{S}^{t-1^{1-\alpha}} \tag{2.9}$$

with $\alpha \in \{0.1, 0.8\} \approx 0$.

From the endogenous coefficients reported in Table 2.7 we notice that they do not change much compared to those reported in column (4) of Table 2.5 which is reassuring that our PML/FP á la Aguirregabiria & Mira (2007) estimates are consistent.

Table 2.7: Estimation of endogenous effects J_y with asymmetric influence by NPL–Λ algorithm

vars	**NPL–Λ**	
	$\alpha = .1$	$\alpha = .8$
	(1)	(2)
unemployed	3.099***	3.057***
	(0.731)	(0.267)
professional	-0.796	-0.976
	(0.624)	(0.650)

Continued on next page

	(1)	(2)
domestic	-1.545	-0.669
	(1.075)	(1.220)
commercial	-3.359**	-3.429**
	(1.062)	(1.143)
industrial	3.519***	3.639***
	(0.257)	(0.265)
log-like	-154100	-154060
obs	165114	
x_i	yes	yes
$w_p X_p$	yes	yes
$u_{p,y}$	yes	yes

Notes: Standard errors in parentheses. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. Beliefs unemployed $s_{p,0}$; professional $s_{p,1}$; domestic $s_{p,2}$; commercial $s_{p,3}$; industrial $s_{p,4}$.

2.4.3 Evidence for identifying assumptions

We now turn present evidence showing that our identifying assumptions are likely to hold. Like most researchers, we are working under the assumption that we have a good measure of social group. To justify our use of ecclesiastical parish boundaries, we have already provided anecdotal evidence suggesting that social networks were "local" (i.e. geography mattered) and ecclesiastical parishes played a major role in the community.

To motivate assumptions **A.**2.2.3 we first provide evidence corroborating that ecclesiastical parishes within a BW were similar. We show that the 1855 Metropolis Management Act, that merged ecclesiastical parishes into BW, created visible differences between BW. For this purpose, we use information on parish receipts and rates. Additionally, we test whether the characteristics of individuals living at the border of two neighbouring parishes within the same BW were significantly similar, which shouldn't be the case if there is sorting at that lower geographic level. We finally use preliminary rent information collected at the street level to show that it is reasonable to think that individuals were "as if" randomly allocated within a parish.

Parish receipts and rates within BW

Given that parishes were allowed to tax their members while providing relief to the paupers it naturally led to affluent parishes being unwilling to accept anyone who could become a charge on the local finances. Initially, mobility restrictions were established dictating responsibility for the poor to their birth parish or to the parish where they had lived for the past three years. A series of acts were later enacted so that the financial burden of paupers was shared on a union-wide basis rather than a parish-wide basis.[33] Therefore there was free mobility within a BW.

Figures 2.5 and 2.6 show the total receipts per inhabitant and the accessible value per inhabitant by BW. There appears to be substantial differences between neighbouring BW while none for parishes within the same BW. Given the fact that this information was public (we found it in a published article of the Economist in 1883), it is reasonable to assume BW boundaries were intimately known by its residents, especially the poor.(Snell 2009) From the local tax receipts in 1881[34] we see wide variation in the wealth of administrative areas. Taken together these evidences suggest that there was a lot of variation in terms of wealth across BW which would have been noticeable to residents when choosing their location. With BW-wide rates, location decisions should have primarily been based on this geographical unit.

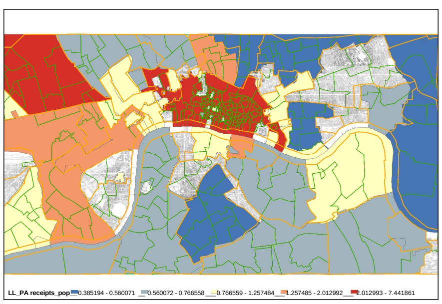


Figure 2.5: Total receipts (in £) per inhabitant by administrative area (BW/Vestries)

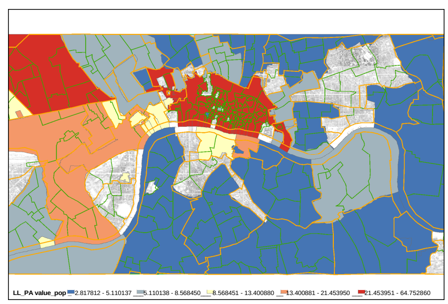


Figure 2.6: Assessable value (in £) per inhabitant by administrative area (BW/Vestries)

Figure 2.7: Total receipts

Source: The Economist Newspaper Ltd, London (1883)

[33]"First, there was the law which made the poor "irremoveable" poor on the common fund of the union, instead of on the parochial rates. Then there was the Union Assessment Act, which equalised the principle of assessment to the common fund throughout the several parishes of the same union, making them contribute to the common fund in proportion to the rateable value for the property in the parish, instead of in proportion to their own previous parochial contributions. Then there was the great reform of last year (1865), the Union Chargeability Act, which changed all the poor on the common fund, so abolishing the temptation offered to close parishes to keep out the poor, unless they could also keep them out of the union itself, - and this is rarely practically possible, - which rendered the poor irremoveable after a single year's residence, instead of three years, and which gave the power of removal to the more intelligent union guardians instead of the less intelligent parish overseers". [The Economist, 1866]

[34]see also figure 2.12 in appendix 2.6.1.

Sorting or not sorting at the parish level

Simple models of residential choice suggest that if parishes boundaries are important determinants of labour market outcomes and individuals know and care about this, there should be substantial sorting along these ecclesiastical boundaries. Households should thus be willing to pay more to live in a "better" parish even if houses and neighbourhoods are very similar on either side of a parish border. Sorting at the ecclesiastical parish level will bias estimates toward finding a positive association between parish quality and employment rate, unless one fully controls for these other differences across boundaries.

However, as explained by the epigraph taken from the Economist in 1857, there was an increasing tendency, during the years succeeding the 1855 Management Act, of residents to sort themselves into locations not based any more on the division of labour but rather a "disposition to associate with equals" based on wealth.[35] Given that parishes within the same BW were facing the same tax burden and were subject to similar redistribution policies we may argue such BW boundaries were the relevant units at which "class–colonies" were emerging.

Additionally, we construct a test to see whether there is any "at-the-border" correlation in unobservables among residents (i.e. at the common border level between two neighbouring parishes, τ_β level), after taking into account the selection based on the BW level (i.e. controlling for τ_b). Given the impossibility to use unobservables to construct such a test we use instead some observables characteristics obtained from the census data (i.e. sex composition, number of children in the household, and of servants as a proxy for wealth, percentage of married couples and share of individuals that have migrated). Conceptually, this methodology is equivalent to testing whether differences in means of exogenous characteristics on opposite sides of social boundaries are statistically zero.

Consider the set of all BW borders as $\mathcal{B}$. Let us define $\beta(p, p') \in \mathcal{B}$ as a border between an ecclesiastical parish p and p' belonging to the same BW. Define a buffer h to this border β and call $\beta_h = \{i \in I \mid d(l_i, \beta) \leq h\}$ as the set of all individuals i that live in location l_i within distance h to a point in the shared border β.

Define Z_p as the random variable Z for individuals residing in ecclesiastical parish p once we have controlled for the BW to which ecclesiastical parish p belongs to using a fixed effect linear regression. Similarly, define $Z_{p'}$ as the random variable Z for individuals belonging to "control" parish of p (i.e. adjacent ecclesiastical parish of p). Now, for a given distance r denote $\beta_r = \{i \in I \setminus \beta_h \mid d(l_i, \beta_h) \leq r\}$ as those observations

[35] "If we secretly regard wealth as the measure of importance, we are awkward in different ways with those richer or poorer than ourselves" *The Economist, June 20, 1857 (p. 670).*

in location l_i that are no more that r-meters apart from any individual belonging to buffer h-meters from border β. Our identifying assumption A.2.2.3 translates in this setting to

$$\lim_{r\to 0} Corr_\beta(\mathbb{E}_{\cdot|\beta_h}[Z_p], \mathbb{E}_{\cdot|\beta_r}[Z_p]) = \lim_{r\to 0} Corr_\beta(\mathbb{E}_{\cdot|\beta_h}[Z_p], \mathbb{E}_{\cdot|\beta_r}[Z_{p'}]) \tag{2.10}$$

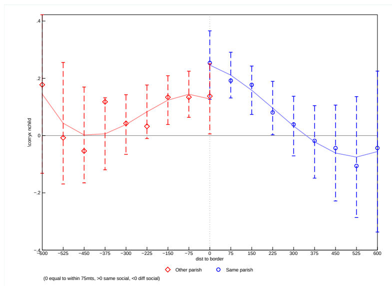
If Z behaves as a random variable at the border β then condition (2.10) should hold. In contrast, if Z responds differently at either side of the border then condition (2.10) is no longer required to hold. In fact, if there is sorting patterns along parishes. exogenous variables for individuals sharing the same social group should be more strongly correlated than for individuals belonging to different social groups. We would expect the correlation in the characteristics between individuals residing within the buffer zone and those living outside it, while still belonging to the same social group, to be larger than the correlation with those equally close but belonging to a different social group. In brief, we should see no discontinuous jump in those observables characteristics that are potentially exogenous (such as age and sex composition).

The following figure 2.8 depict evidence for a buffer $h = 40$ *mts* and bins (r) of 75 *mts*. The horizontal axis varies the distance to buffer observations with positive values reserved to those individuals belonging to the same parish while negative values depict individuals belonging to "control" parishes. On the vertical axis we plot the corresponding correlation. Our identifying assumption imply that there should not be a discontinuity at the origin if one compares the correlation among neighbours of the same social parish and the correlation among neighbours of the control social parish.[36]

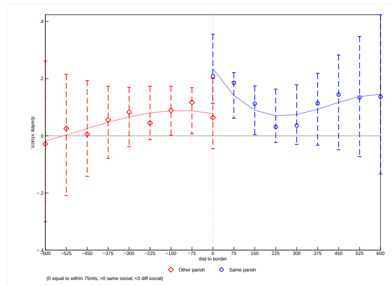
In figure 2.8 we see that none of the exogenous variables, apart from the share of married head of households, exhibit any discontinuities in their correlations while distinguishing by actual parish and neighbouring ones. Social interactions could be an important determinant for the marriage market, similarly to the labour market. Therefore, the discontinuity found in the share of married head of households should not be surprising. We conclude that there is evidence of no selection at the ecclesiastical parish boundaries (i.e social group borders) once BW fixed effects are accounted for.

We have also compiled some tentative evidence on rents at the building level (Stewart 1900). Our preliminary data (see appendix 2.6.1) shows that within–BW variation is lower than between–BW which suggests that differences in house rents for parishes within the same BW were lower than differences for parishes across different BW. However more work on this is needed.

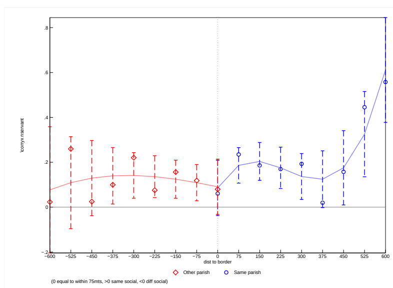
[36]We show the results for a buffer of = 40 *mts* and a degree third polynomial, but the results are robust to different buffers and polynomial degrees (see Appendix 2.6).



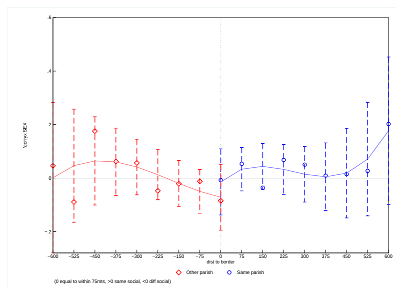
(a) number of children



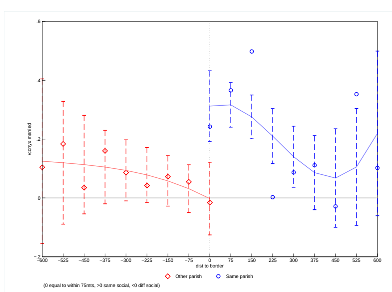
(b) parish birth=residence



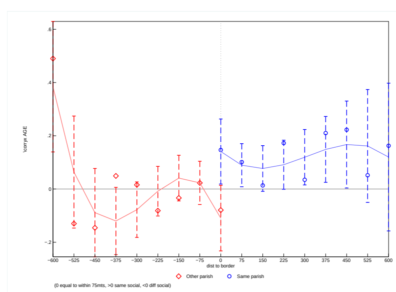
(c) number of servants



(d) male



(e) married



(f) age

Figure 2.8: Correlation in characteristics for $h = 40$ mts varying neighbours at 75 mts bins, polynomial degree 3

2.5 Conclusion

In this chapter we have presented the identification and estimation of a multinomial choice model with social interactions and asymmetric influence. The model allows for correlated effects at the group level and includes a spatial weighting matrix to capture potential interactions and/or strength of ties. We establish the identification of the endogenous and exogenous interactions when there is enough variation on the behavioural influences within a group. This extends prior work on social interactions focusing on binary outcomes with asymmetric influence (Lee et al. 2014) and multinomial choice model with symmetric influence (Brock & Durlauf 2006). It also depicts how variation across and within groups may be exploited for identification with non–network data (Bramoullé 2013, Goldsmith-Pinkham & Imbens 2013). We use a recursive pseudo maximum likelihood estimation with equilibrium fixed point subroutine (Aguirregabiria & Mira 2007) to provide consistent and asymptotically efficient estimates of our structural parameters. The empirical framework developed in this chapter may be applied to other areas involving local interactions and categorical outcomes such as criminal activities, modes of transport, or technology adoption.

As an empirical application, we examine how social groups affect occupational choices in Victorian London. We construct a new dataset which allows the geographical localisation of the 1881 full census data. We define social groups using the ecclesiastical parish boundaries and exploit a two–tier administrative system to deal with self–selection into groups. We argue that ecclesiastical parishes were a defining feature of social networks and individuals' location decisions were based on BW, providers of public good services. Our results indicate that social parishes play a role in determining labour market outcomes among Londoners in 1881. Once multiple equilibria in the consistent beliefs constraint and group unobservables are accounted for, an increase in the share of a industrial occupation in one's parish peers increases one's own probability of being employed at that same occupation, while for commercial occupations peer's competition is predominant. We also report that a higher expected incidence of unemployed peers leads to a larger likelihood of being unemployed. Social interactions do not seem to matter for occupational choice at both ends of skills' distribution (i.e. for domestic and professional occupations).

While our specific data allows us to investigate a historical period, our results might be relevant for social network effects in contemporary studies. In the modern world of easy mobility and technological information, we content that geography–related measures could capture the most relevant features of social networks. In the 19th century such measure had more relevant content than nowadays. Moreover, the religious dimension of our measure offers a plausible additional dimension given that church attendance remained mandatory as a legacy of the Tudor era.

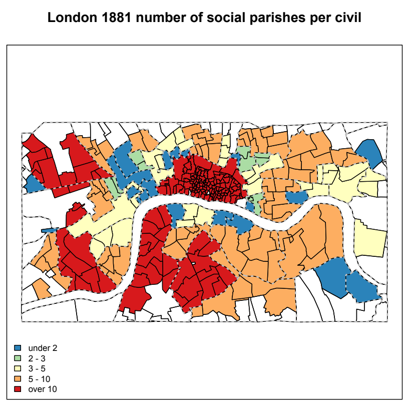
Relying on our historical period also enables us to circumvent the self–selection into social group problems thanks to the curious form of local federalism based on a two-tier administrative system present at that time. We exploit the fact that public goods were provided at a higher tier and consequently determined location decision while community identity were still largely determined at a more local level.

Our chapter helps us understand how social networks play a role in labour market outcomes above and beyond neighbourhood (Topa & Zenou 2014). While most studies have looked employment status, our study documents spatial clustering in occupation within a city and can shed light on how otherwise homogeneous societies may differ substantially due to the composition of their social reference group. A strong endogenous effect suggests that any program that targets employment in particular sectors, will have a spillover effect: increasing the employment likelihood of someone else in the network. We show that failing to account for asymmetric influence and ignoring possible correlated effects may bias the endogenous effect on occupational choices.

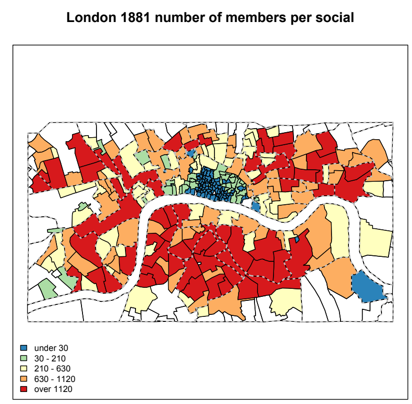
Studying social interactions and labour market outcomes can also help us to understand social trends and transformations such as social mobility and industrialisation. Inter-generational occupational and spatial mobility may remain low because workers seek to use their inherited social connections to find jobs more easily as documented for instance by Borjas (1994) and Munshi & Wilson (2008). This is a interesting question for future work.

2.6 Appendix

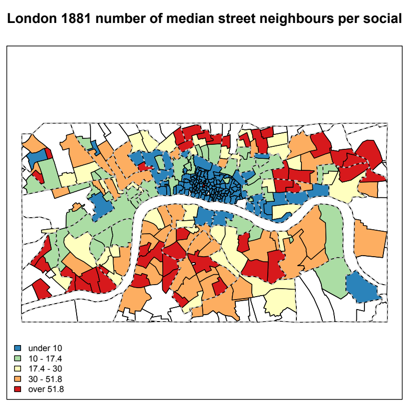
2.6.1 Further descriptives



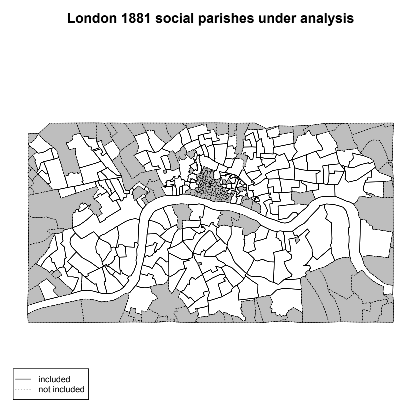
(a) $\mid P \mid$



(b) n_p

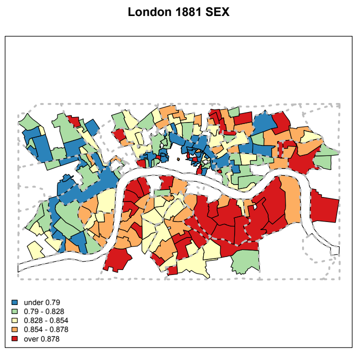


(c) nei_p

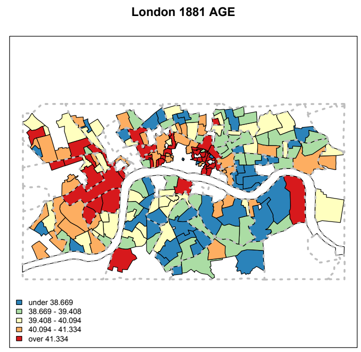


(d) Area under investigation

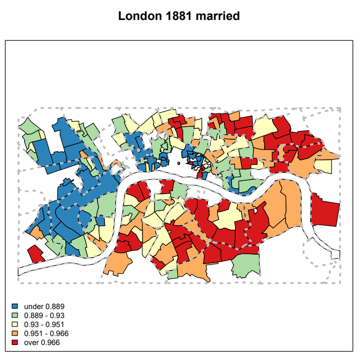
Figure 2.9: Social and civil parish: variation in $\mid P(b) \mid, n_p, nei(i)$



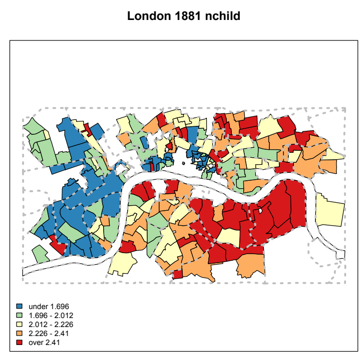
(a) Percentage of men per parish



(b) Average age per parish



(c) Percentage of married individual per parish



(d) Average number of children per parish

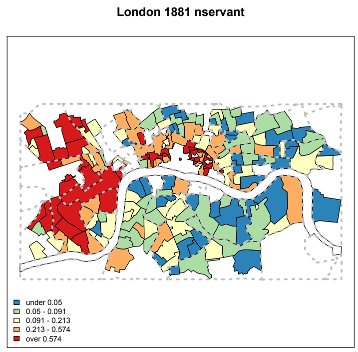
2.6.2 CML estimation

The following is the generalization to the multinomial logit case for the CML to difference out fixed effect at the group level (Chamberlain 1980, Gabrielsen 1978).

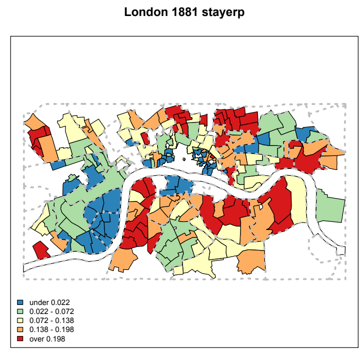
One will need to restrict the sample N to reference groups where there is variation in terms of occupational choices (which we denote N'). Define $\mu_{p,iy} = 1$ if $y_{p,i} = y$, $\mu_{p,iy} = 0$ otherwise. The probability distribution of the restricted sample conditioning on $t_{p,y} = \sum_{i \in p} \mu_{p,iy}$ for every y, which is a sufficient statistic for every $u_{p,y}$, leads to the following Conditional Maximum Likelihood function for the sub–sample N'

$$\mathcal{L}_{N'} = \frac{1}{N'} \sum_{p \in P} \sum_{i \in p} \log \frac{\exp\left(\beta' \sum_{i,y} z_{p,iy} \mu_{p,iy}\right)}{\sum_{\lambda \in \Lambda_p} \exp\left(\beta' \sum_{i,y} z_{p,iy} \lambda_{iy}\right)}. \tag{2.11}$$

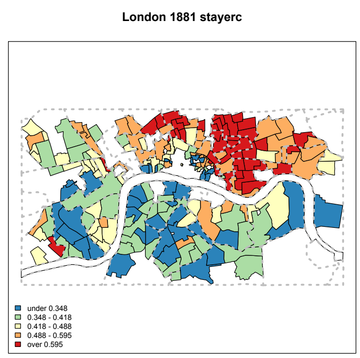
Where $\Lambda_p = \Big\{ \lambda = (\lambda_{1,0}, \cdots, \lambda_{n_p,L-1}) \mid \lambda_{iy} = 0 \text{ or } 1, \sum_y \lambda_{iy} = 1, \sum_{i \in p} \lambda_{iy} = t_{py}, y = 1, \cdots,$



(e) Number of servants



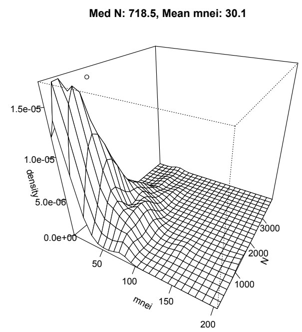
(f) Percentage of residents in parish of birth



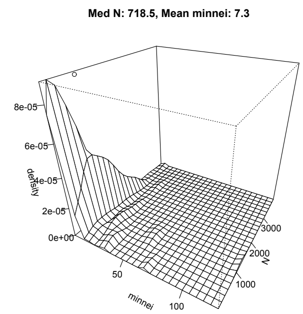
(g) Percentage of resident in county of birth

Figure 2.10: Descriptives by ecclesiastical parish

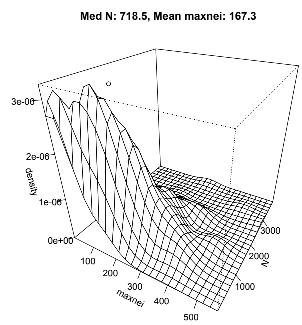
$L - 1\}$



(a) median *nei*



(b) min *nei*



(c) max *nei*

Figure 2.11: Empirical distribution of parish members against statistics on street neighbours, London 1881

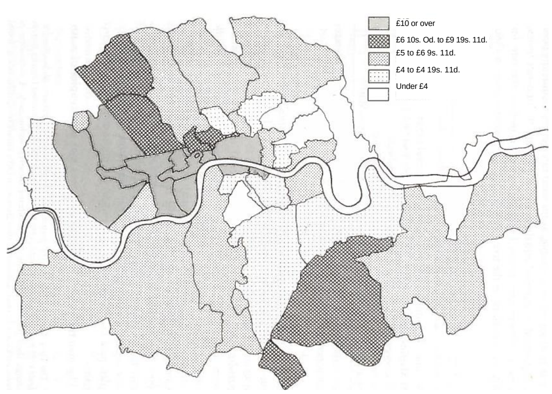


Figure 2.12: Local Areas: Rateable Value per Head, 1881. Davis (1988)

Table 2.8: Balanced Sample across merged observations in non institutional dwellings

	Not merged			Merged			$H_0 : 1-2=0$	
	N	Mean (1)	SE	N	Mean (2)	SE	t-stat	pval
All individuals								
Male	802,735	0.450	0.001	895,712	0.473	0.001	30.430	0.000
Age	802,746	32.521	0.014	895,718	32.767	0.013	13.104	0.000
Pop Age 25-34	802,746	0.279	0.000	895,718	0.276	0.000	-4.087	0.000
Pop Age 35-44	802,746	0.207	0.000	895,718	0.205	0.000	-3.282	0.001
Pop Age 45-60	802,746	0.194	0.000	895,718	0.203	0.000	14.673	0.000
Native	802,746	0.934	0.000	895,718	0.920	0.000	-32.871	0.000
Labour Force	802,202	0.665	0.001	894,880	0.694	0.000	40.229	0.000
Married	799,622	0.590	0.001	892,655	0.608	0.001	24.750	0.000
Individuals in non extreme border parishes†								
Male	150,171	0.473	0.001	491,167	0.467	0.001	-3.737	0.000
Age	150,171	32.822	0.032	491,170	32.731	0.018	-2.499	0.012
Pop Age 15-24	150,171	0.319	0.001	491,170	0.320	0.001	0.574	0.566
Pop Age 25-34	150,171	0.266	0.001	491,170	0.273	0.001	5.446	0.000
Pop Age 35-44	150,171	0.208	0.001	491,170	0.202	0.001	-4.605	0.000
Pop Age 45-60	150,171	0.207	0.001	491,170	0.205	0.001	-2.065	0.039
Native	150,171	0.904	0.001	491,170	0.901	0.000	-2.316	0.021
Labour Force	150,017	0.700	0.001	490,731	0.708	0.001	6.204	0.000
Married	149,572	0.600	0.001	489,443	0.590	0.001	-6.532	0.000
Individuals in non border parishes‡								
Male	97,259	0.489	0.002	293,892	0.488	0.001	-0.523	0.601
Age	97,259	32.835	0.040	293,894	32.761	0.023	-1.602	0.109
Pop Age 15-24	97,259	0.318	0.001	293,894	0.322	0.001	1.968	0.049
Pop Age 25-34	97,259	0.264	0.001	293,894	0.267	0.001	1.532	0.126
Pop Age 35-44	97,259	0.210	0.001	293,894	0.204	0.001	-3.895	0.000
Pop Age 45-60	97,259	0.207	0.001	293,894	0.207	0.001	-0.052	0.959
Native	97,259	0.896	0.001	293,894	0.891	0.001	-3.615	0.000
Labour Force	97,160	0.701	0.001	293,560	0.703	0.001	1.211	0.226
Married	96,866	0.615	0.002	292,865	0.616	0.001	0.863	0.388

‡ *border parishes* are: Battersea, Bow, Bromley St Leonard, Brompton, Camberwell, St Dunstan Stepney/Mile End, Mile End New Town, Poplar, St George Hanover Square, St James Clerkenwell, St Leonard Shoreditch, St Luke Chelsea, St Luke Old Street, St Margaret Westminster, St Mary Abbots Kensington, St Mary Lambeth, St Mary Paddington, St Mary Rotherhithe, St Marylebone, St Matthew Bethnal Green, St Nicholas Deptford, St Pancras, St Paul Deptford, Greenwich, St Anne Kensington, Brompton, Islington. † extreme border parishes are: *border parishes* minus St George Hanover Square, St Luke Old Street, St Dunstan Stepney/Mile End, Mile End New Town, St Marylebone.

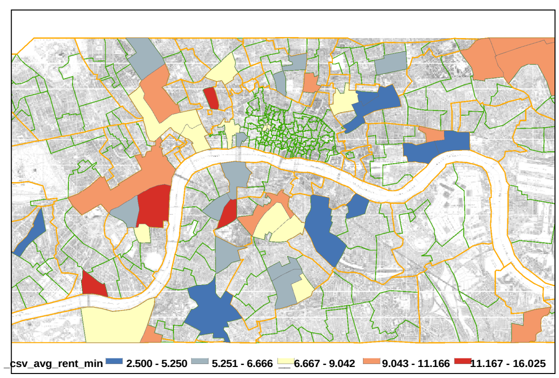


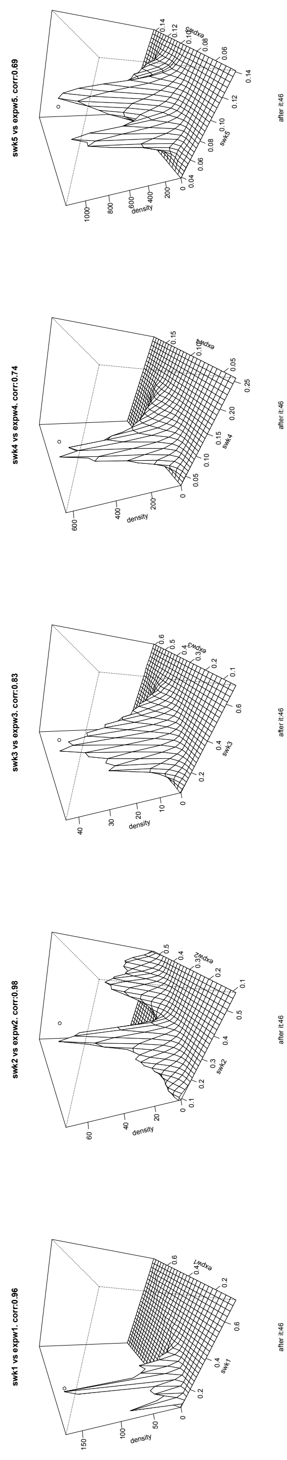
Figure 2.13: Avg. rent per room (in £, lower bound) by ecclesiastical parish

2.6.3 Other Montecarlos results

Table 2.9: Montecarlo simulations and estimation when $s_{p,y}$ is observed

	$J0 = 3.3$			$J1 = 2.5$			$J2 = 2$			$J3 = 3.2$			$J4 = 3.6$			$\min nei_i$ n_p	
$\mid P \mid$	mean	med	sd	mean	med	sd	mean	med	sd	mean	med	sd	mean	med	sd	mean	
								ML when $(s_{p,y}, w)$ are observed									
5	3.44	3.18	3.15	2.75	2.45	1.98	1.92	1.95	1.28	2.75	2.50	7.18	5.81	4.92	13.31	8.2	4471.0
mse		*9.86*			*3.94*			*1.63*			*51.31*			*180.36*			
10	3.47	3.43	0.99	2.59	2.55	0.95	1.94	1.86	0.84	3.30	3.25	3.31	4.69	4.01	7.96	7.3	8478.6
mse		*0.99*			*0.90*			*0.69*			*10.87*			*63.88*			
50*	3.31	3.35	0.27	2.52	2.57	0.39	2.01	2.00	0.32	3.46	3.32	1.15	3.86	4.00	2.60	7.3	42994.5
mse		*0.07*			*0.15*			*0.10*			*1.37*			*6.75*			
100**	3.32	3.31	0.20	2.47	2.46	0.30	2.02	2.00	0.25	3.37	3.52	0.73	3.58	3.58	1.73	7.8	85752.1
mse		*0.04*			*0.09*			*0.06*			*0.55*			*2.98*			

* 105 simulations, ** 107 simulations



(a) $cor(w_p s_{p,0}, w_p^* \hat{s}_{p,0})$ (b) $cor(w_p s_{p,1}, w_p^* \hat{s}_{p,1})$ (c) $cor(w_p s_{p,2}, w_p^* \hat{s}_{p,2})$ (d) $cor(w_p s_{p,3}, w_p^* \hat{s}_{p,3})$ (e) $cor(w_p s_{p,4}, w_p^* \hat{s}_{p,4})$

Figure 2.14: Correlation between $w_p s_{p,y}$ and $w_p^* \hat{s}_{p,y}$

Table 2.10: Montecarlo simulations and estimation of d when $s_{p,y}$ is not observed

	$d_1 - d_0 = 0.2$			$d_2 - d_0 = 0.5$			$d_3 - d_0 = 0.1$			$d_4 - d_0 = 0.4$		
$\mid P \mid$	mean	med	sd	mean	med	sd	mean	med	sd	mean	med	sd
A. PML/FP, $s_{p,y}$ not observed, w observed												
5	0.22	0.22	0.18	0.57	0.53	0.29	0.04	0.08	0.42	0.45	0.46	0.29
mse		*0.03*			*0.09*			*0.17*			*0.09*	
10	0.19	0.19	0.10	0.54	0.52	0.17	0.07	0.08	0.21	0.37	0.42	0.19
mse		*0.01*			*0.03*			*0.04*			*0.04*	
50*	0.20	0.20	0.03	0.51	0.50	0.07	0.11	0.10	0.07	0.39	0.40	0.07
mse		*0.00*			*0.01*			*0.00*			*0.01*	
100**	0.20	0.20	0.02	0.51	0.51	0.04	0.10	0.10	0.05	0.39	0.39	0.04
mse		*0.00*			*0.00*			*0.00*			*0.00*	
B. PML/FP, $s_{p,y}$ not observed, w not observed instead w^* is observed												
5	0.20	0.17	0.19	0.96	0.94	0.26	0.01	-0.01	0.32	0.41	0.42	0.36
mse		*0.04*			*0.28*			*0.11*			*0.13*	
10	0.16	0.16	0.10	0.88	0.87	0.15	-0.05	-0.05	0.18	0.34	0.34	0.16
mse		*0.01*			*0.17*			*0.06*			*0.03*	
50*	0.17	0.16	0.04	0.87	0.87	0.07	-0.06	-0.05	0.07	0.34	0.34	0.06
mse		*0.00*			*0.14*			*0.03*			*0.01*	
100**	0.15	0.15	0.03	0.85	0.85	0.05	-0.09	-0.08	0.05	0.33	0.32	0.04
mse		*0.13*			*0.04*			*0.01*			*0.01*	

* 105 simulations, ** 107 simulations. *med*: median, *sd*: standard deviation, *mse*: mean square error.

Table 2.11: Montecarlo simulations and estimation of a when $s_{p,y}$ is not observed

	$a_1 - a_0 = -0.1$			$a_2 - a_0 = 0.6$			$a_3 - a_0 = -0.6$			$a_4 - a_0 = -0.4$		
$\mid P \mid$	mean	med	sd	mean	med	sd	mean	med	sd	mean	med	sd
A. PML/FP, $s_{p,y}$ not observed, w observed												
5	-0.10	-0.10	0.05	0.61	0.60	0.07	-0.61	-0.61	0.08	-0.40	-0.40	0.06
mse		*0.00*			*0.00*			*0.01*			*0.00*	
10	-0.10	-0.10	0.03	0.61	0.61	0.03	-0.61	-0.60	0.04	-0.40	-0.39	0.04
mse		*0.00*			*0.00*			*0.00*			*0.00*	
50*	-0.10	-0.10	0.01	0.60	0.60	0.01	-0.60	-0.60	0.02	-0.40	-0.40	0.02
mse		*0.00*			*0.00*			*0.00*			*0.00*	
100**	-0.10	-0.10	0.01	0.60	0.60	0.01	-0.60	-0.60	0.01	-0.40	-0.40	0.01
mse		*0.00*			*0.00*			*0.00*			*0.00*	
B. PML/FP, $s_{p,y}$ not observed, w not observed instead w^* is observed												
5	-0.10	-0.09	0.05	0.63	0.62	0.07	-0.61	-0.61	0.08	-0.39	-0.40	0.06
mse		*0.00*			*0.01*			*0.01*			*0.00*	
10	-0.10	-0.10	0.03	0.62	0.62	0.03	-0.61	-0.60	0.04	-0.40	-0.39	0.04
mse		*0.00*			*0.00*			*0.00*			*0.00*	
50*	-0.10	-0.10	0.01	0.62	0.61	0.01	-0.61	-0.61	0.02	-0.40	-0.40	0.02
mse		*0.00*			*0.00*			*0.00*			*0.00*	
100**	-0.10	-0.10	0.01	0.62	0.61	0.01	-0.61	-0.61	0.01	-0.40	-0.40	0.01
mse		*0.00*			*0.00*			*0.00*			*0.00*	

* 105 simulations, ** 107 simulations. *med*: median, *sd*: standard deviation, *mse*: mean square error.

Table 2.12: Montecarlo simulations and estimation of $J_y \geq 3.2$ for all $y \in \Omega$ when $s_{p,y}$ is not observed

	$J0 = 3.5$			$J1 = 3.3$			$J2 = 3.2$			$J3 = 3.4$			$J4 = 3.8$			$\min nei_i$	n_p
$\mid P \mid$	mean	med	sd	mean	med	sd	mean	med	sd	mean	med	sd	mean	med	sd	mean	
							A. PML/FP, $s_{p,y}$ not observed, w observed										
5	2.86	3.53	2.67	3.25	3.28	0.39	3.20	3.21	0.33	2.63	3.25	2.66	2.36	3.90	6.14	9.30	4411.32
mse		*7.47*			*0.15*			*0.11*			*7.62*			*39.45*			
10	3.01	3.32	1.14	3.27	3.31	0.26	3.19	3.19	0.21	2.85	3.24	1.92	2.36	3.26	4.41	7.31	8369.72
mse		*1.53*			*0.07*			*0.05*			*3.96*			*21.30*			
50*	3.48	3.46	0.19	3.31	3.32	0.06	3.19	3.18	0.09	3.21	3.27	0.58	3.53	3.61	1.90	7.76	42964.68
mse		*0.04*			*0.00*			*0.01*			*0.37*			*3.66*			
100**	3.48	3.48	0.13	3.31	3.31	0.04	3.18	3.18	0.06	3.32	3.31	0.33	3.45	3.57	1.19	7.73	88280.75
mse		*0.02*			*0.00*			*0.00*			*0.12*			*1.53*			
							B. PML/FP, $s_{p,y}$ not observed, w not observed instead w^* is observed										
5	1.99	3.09	3.03	3.11	3.34	0.92	1.68	1.79	0.75	0.68	0.83	2.33	0.09	0.88	5.42	9.30	4411.32
mse		*11.38*			*0.87*			*2.86*			*12.73*			*42.79*			
10	1.99	2.60	2.41	3.32	3.39	0.42	1.82	1.86	0.49	0.90	1.20	1.74	-0.39	-0.26	3.50	7.31	8369.72
mse		*8.03*			*0.18*			*2.15*			*9.28*			*29.65*			
50*	3.18	3.33	0.59	3.37	3.37	0.10	2.00	2.01	0.25	1.24	1.26	0.86	0.14	0.33	1.98	7.76	42964.68
mse		*0.44*			*0.02*			*1.50*			*5.42*			*17.25*			
100**	3.32	3.40	0.45	3.37	3.39	0.09	2.04	2.04	0.18	1.27	1.30	0.59	-0.19	-0.11	1.32	7.73	88280.75
mse		*0.24*			*0.01*			*1.39*			*4.90*			*17.61*			
choices		*0.08*			*0.46*			*0.36*			*0.05*			*0.04*			

* 116 simulations, ** 88 simulations. *med*: median, *sd*: standard deviation, *mse*: mean square error.

Table 2.13: Montecarlo simulations and estimation of J_y when: A. $J_y < 3.2$ for all $y \in \Omega$; B. $\exists y' \in \Omega, J_{y'} < 0$, when $s_{p,y}$ is not observed

	$J0$			$J1$			$J2$			$J3$			$J4$			$\min nei_i$	n_p
$\mid P \mid$	mean	med	sd	mean	med	sd	mean	med	sd	mean	med	sd	mean	med	sd	mean	
A. PML/FP, $s_{p,y}$ not observed, w observed																	
i) $J_y < 3.2$		$J0 = 2.8$			$J1 = 2$			$J2 = 1.5$			$J3 = 2.7$			$J4 = 3.1$			
50†	2.77	2.81	0.31	1.89	1.92	0.52	1.49	1.50	0.37	2.32	2.57	1.03	2.53	3.07	2.58	7.41	42629.15
mse		*0.10*			*0.28*			*0.14*			*1.20*			*6.91*			
100‡	2.76	2.79	0.21	1.93	1.93	0.34	1.51	1.49	0.23	2.53	2.59	0.61	2.56	2.87	1.84	7.28	84779.00
mse		*0.05*			*0.12*			*0.05*			*0.40*			*3.64*			
ii) $J_{y'} < 0$		$J0 = 3.3$			$J1 = -3.1$			$J2 = 2$			$J3 = 3.2$			$J4 = 3.6$			
50*	3.29	3.29	0.10	-3.18	-2.86	2.38	1.98	1.98	0.19	3.13	3.23	0.48	3.31	3.51	1.17	7.80	43699.40
mse		*0.01*			*5.62*			*0.04*			*0.23*			*1.44*			
100**	3.29	3.30	0.07	-3.02	-2.90	1.61	1.99	2.01	0.14	3.09	3.15	0.33	3.44	3.44	0.68	7.55	86268.57
mse		*0.01*			*2.57*			*0.02*			*0.12*			*0.48*			
B. PML/FP, $s_{p,y}$ not observed, w not observed instead w^* is observed																	
i) $J_y < 3.2$		$J0 = 2.8$			$J1 = 2$			$J2 = 1.5$			$J3 = 2.7$			$J4 = 3.1$			
50†	2.70	2.73	0.32	1.67	1.69	0.52	0.13	0.13	0.23	0.82	0.88	0.62	0.89	0.89	1.45	7.41	42629.15
mse		*0.11*			*0.38*			*1.94*			*3.90*			*6.97*			
100‡	2.70	2.76	0.23	1.69	1.73	0.36	0.15	0.14	0.17	0.81	0.80	0.47	0.98	0.99	1.05	7.28	84779.00

Continued on next page

		$J0$			$J1$			$J2$			$J3$			$J4$		$\min nei_i$	n_p
mse		*0.06*			*0.22*			*1.86*			*3.78*			*5.59*			
ii) $J_{y'} < 0$		$J0 = 3.3$			$J1 = -3.1$			$J2 = 2$			$J3 = 3.2$			$J4 = 3.6$			
50*	3.38	3.39	0.12	-2.51	-2.27	1.94	0.67	0.66	0.22	1.20	1.22	0.57	1.09	1.17	1.15	7.80	43699.40
mse		*0.02*			*4.08*			*1.81*			*4.31*			*7.63*			
100**	3.39	3.39	0.08	-2.44	-2.43	1.33	0.68	0.67	0.16	1.17	1.18	0.39	1.18	1.21	0.76	7.55	86268.57
mse		*0.02*			*2.19*			*1.75*			*4.29*			*6.45*			
choice $J_y < 3.2$		*0.25*			*0.28*			*0.28*			*0.10*			*0.10*			
choice $J_{y'} < 0$		*0.51*			*0.08*			*0.26*			*0.08*			*0.07*			

† 130 sim, ‡ 126 sim, * 132 simulations, ** 99 simulations. *med*: median, *sd*: standard deviation, *mse*: mean square error.

2.6.4 Individual and contextual effects

Table 2.14: Exogenous (k_y, c_y) and contextual d_y effects for Homogeneous and Heterogeneous

Beliefs	**Homogeneous**				**Heterogeneous**			
vars	Naive estimation			PML/FP	Naive estimation		PML/FP	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
k_1	-0.049	1.174	0.433	4.298**	1.026*	-0.042	1.568**	0.114
	(0.086)	(1.293)	(1.544)	(1.583)	(0.483)	(0.915)	(0.499)	(0.921)
k_2	1.225***	4.152***	3.764**	0.533	1.206**	0.55	0.778	2.973***
	(0.076)	(1.166)	(1.367)	(1.38)	(0.444)	(0.779)	(0.48)	(0.87)
k_3	-2.174***	0.298	0.161	0.872	-1.005*	-1.787*	-1.082*	-1.543·
	(0.112)	(1.194)	(1.408)	(1.408)	(0.449)	(0.836)	(0.466)	(0.855)
k_4	1.305***	4.286***	3.422**	3.595**	1.926***	1.108	1.295**	0.779
	(0.084)	(1.066)	(1.267)	(1.289)	(0.398)	(0.716)	(0.415)	(0.74)
male_1	3.334***	3.364***	3.36***	3.352***	3.352***	3.358***	3.333***	3.349***
	(0.042)	(0.042)	(0.042)	(0.042)	(0.043)	(0.043)	(0.042)	(0.042)
male_2	0.774***	0.799***	0.798***	0.812***	0.758***	0.757***	0.813***	0.797***
	(0.032)	(0.032)	(0.032)	(0.032)	(0.033)	(0.033)	(0.032)	(0.033)
male_3	6.37***	6.395***	6.415***	6.409***	6.432***	6.443***	6.416***	6.436***
	(0.085)	(0.085)	(0.085)	(0.085)	(0.085)	(0.085)	(0.085)	(0.085)
male_4	3.324***	3.352***	3.367***	3.357***	3.401***	3.412***	3.374***	3.4***
	(0.027)	(0.027)	(0.027)	(0.027)	(0.028)	(0.028)	(0.027)	(0.028)
age_1	-0.021***	-0.021***	-0.021***	-0.021***	-0.023***	-0.023***	-0.022***	-0.022***
	(0.001)	(0.002)	(0.002)	(0.001)	(0.002)	(0.002)	(0.002)	(0.002)

Continued on next page

vars	Homo, Naive			Homo, PML/FP	Hete, Naive		Hete, PML/FP	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
age_2	-0.019***	-0.019***	-0.019***	-0.019***	-0.019***	-0.019***	-0.019***	-0.019***
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
age_3	-0.028***	-0.028***	-0.028***	-0.028***	-0.028***	-0.028***	-0.028***	-0.028***
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
age_4	-0.02***	-0.021***	-0.021***	-0.021***	-0.02***	-0.02***	-0.02***	-0.02***
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
$married_1$	-1.05***	-1.059***	-1.061***	-1.058***	-1.04***	-1.04***	-1.037***	-1.028***
	(0.056)	(0.057)	(0.057)	(0.056)	(0.058)	(0.058)	(0.057)	(0.057)
$married_2$	-0.346***	-0.293***	-0.294***	-0.293***	-0.323***	-0.317***	-0.321***	-0.316***
	(0.05)	(0.05)	(0.05)	(0.05)	(0.051)	(0.052)	(0.051)	(0.051)
$married_3$	-0.623***	-0.597***	-0.602***	-0.599***	-0.682***	-0.678***	-0.657***	-0.655***
	(0.055)	(0.056)	(0.056)	(0.056)	(0.057)	(0.057)	(0.056)	(0.057)
$married_4$	-0.718***	-0.694***	-0.702***	-0.7***	-0.799***	-0.794***	-0.785***	-0.782***
	(0.044)	(0.045)	(0.045)	(0.045)	(0.046)	(0.046)	(0.045)	(0.046)
n $child_1$	-0.032***	-0.027**	-0.027**	-0.027**	-0.023**	-0.023**	-0.025**	-0.024**
	(0.009)	(0.009)	(0.009)	(0.009)	(0.009)	(0.009)	(0.009)	(0.009)
n $child_2$	-0.071***	-0.062***	-0.062***	-0.062***	-0.072***	-0.07***	-0.07***	-0.068***
	(0.008)	(0.008)	(0.008)	(0.008)	(0.009)	(0.009)	(0.008)	(0.009)
n $child_3$	-0.04***	-0.037***	-0.037***	-0.037***	-0.046***	-0.045***	-0.045***	-0.045***
	(0.007)	(0.008)	(0.008)	(0.007)	(0.008)	(0.008)	(0.008)	(0.008)
n $child_4$	-0.001	0.003	0.003	0.002	-0.006	-0.006	-0.007	-0.007
	(0.007)	(0.007)	(0.007)	(0.007)	(0.007)	(0.007)	(0.007)	(0.007)
n $servant_1$	0.022*	0.019·	0.018·	0.015	0.025·	0.025·	0.017	0.024·
	(0.011)	(0.011)	(0.011)	(0.011)	(0.013)	(0.013)	(0.013)	(0.013)

Continued on next page

vars	Homo, Naive			Homo, PML/FP	Hete, Naive		Hete, PML/FP	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
n $servant_2$	-0.27***	-0.308***	-0.304***	-0.299***	-0.289***	-0.29***	-0.274***	-0.283***
	(0.016)	(0.016)	(0.016)	(0.016)	(0.018)	(0.018)	(0.018)	(0.018)
n $servant_3$	-0.324***	-0.344***	-0.343***	-0.343***	-0.252***	-0.254***	-0.251***	-0.254***
	(0.016)	(0.016)	(0.016)	(0.016)	(0.018)	(0.018)	(0.018)	(0.018)
n $servant_4$	-0.176***	-0.191***	-0.189***	-0.189***	-0.067***	-0.069***	-0.068***	-0.068***
	(0.01)	(0.011)	(0.011)	(0.011)	(0.012)	(0.013)	(0.012)	(0.012)
$stayer_1$	-0.415***	-0.384***	-0.385***	-0.383***	-0.325***	-0.326***	-0.333***	-0.333***
	(0.057)	(0.059)	(0.059)	(0.058)	(0.059)	(0.059)	(0.059)	(0.059)
$stayer_2$	-0.032	-0.018	-0.018	-0.019	-0.044	-0.046	-0.042	-0.046
	(0.048)	(0.049)	(0.049)	(0.049)	(0.05)	(0.05)	(0.05)	(0.05)
$stayer_3$	0.074·	0.06	0.061	0.06	0.025	0.024	0.017	0.018
	(0.044)	(0.045)	(0.045)	(0.045)	(0.046)	(0.046)	(0.045)	(0.046)
$stayer_4$	0.294***	0.298***	0.299***	0.297***	0.232***	0.232***	0.223***	0.223***
	(0.039)	(0.04)	(0.04)	(0.04)	(0.041)	(0.041)	(0.041)	(0.041)
mean $male_1$		-6.843***	-5.379***	-2.602*	-2.304***	-1.526***	-0.356	-0.751·
	(0)	(0.64)	(0.894)	(1.071)	(0.278)	(0.296)	(0.365)	(0.386)
mean $male_2$		-3.73***	-1.94*	0.799	-0.1	0.268	1.253***	-0.277
	(0)	(0.585)	(0.807)	(0.989)	(0.255)	(0.271)	(0.354)	(0.428)
mean $male_3$		-6.386***	-4.275***	-2·	-1.959***	-1.271***	0.32	-0.032
	(0)	(0.584)	(0.811)	(1.063)	(0.257)	(0.273)	(0.438)	(0.441)
mean $male_4$		-6.049***	-4.336***	-2.442**	-2.16***	-1.503***	-1.311***	-1.645***
	(0)	(0.526)	(0.738)	(0.933)	(0.233)	(0.248)	(0.322)	(0.338)
mean age_1		0.017	0.024	-0.026	0.013	0.012	0.012	0.017·
	(0)	(0.022)	(0.027)	(0.027)	(0.009)	(0.01)	(0.009)	(0.01)

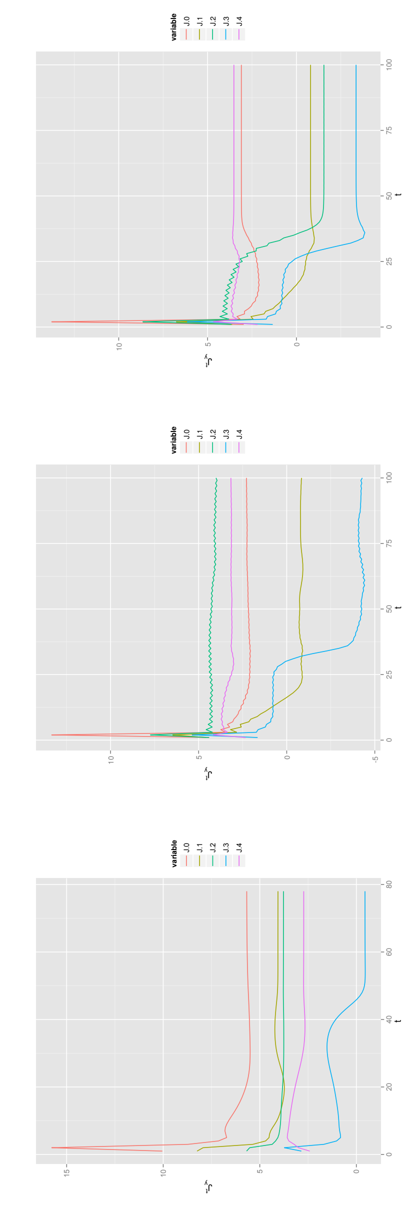
Continued on next page

vars	Homo, Naive			Homo, PML/FP	Hete, Naive		Hete, PML/FP	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
mean age_2		0.024	0.006	0.06*	0.002	-0.008	0.003	-0.009
	(0)	(0.02)	(0.024)	(0.025)	(0.008)	(0.009)	(0.008)	(0.009)
mean age_3		0.023	0.017	-0.011	0.004	-0.002	-0.016·	-0.013
	(0)	(0.02)	(0.024)	(0.025)	(0.008)	(0.009)	(0.008)	(0.009)
mean age_4		0.013	0.017	0.005	0.007	0.002	0	-0.001
	(0)	(0.018)	(0.022)	(0.022)	(0.007)	(0.008)	(0.007)	(0.008)
mean $married_1$		4.453***	3.812***	-0.513	1.117***	0.8*	-0.552·	0.023
	(0)	(0.723)	(0.854)	(1.075)	(0.301)	(0.325)	(0.313)	(0.322)
mean $married_2$		-0.301	-1.24	-1.849·	0.089	0.369	-0.479·	-0.393
	(0)	(0.687)	(0.823)	(0.968)	(0.285)	(0.303)	(0.278)	(0.297)
mean $married_3$		1.607*	0.76	-0.607	0.494·	0.59·	0.41	0.657*
	(0)	(0.732)	(0.878)	(1.073)	(0.288)	(0.308)	(0.289)	(0.303)
mean $married_4$		1.429*	0.556	-1.325	0.529*	0.64*	0.185	0.562*
	(0)	(0.652)	(0.795)	(0.951)	(0.26)	(0.277)	(0.26)	(0.273)
mean n $child_1$		-0.066	-0.066	0.071	-0.092*	-0.095·	-0.103*	-0.143**
	(0)	(0.084)	(0.118)	(0.117)	(0.045)	(0.049)	(0.047)	(0.051)
mean n $child_2$		-0.117	0.167	-0.034	0.231***	0.284***	0.212***	0.303***
	(0)	(0.078)	(0.108)	(0.111)	(0.04)	(0.043)	(0.04)	(0.044)
mean n $child_3$		0.29***	0.104	0.235*	0.203***	0.221***	0.226***	0.214***
	(0)	(0.073)	(0.103)	(0.107)	(0.038)	(0.042)	(0.039)	(0.042)
mean n $child_4$		0.195**	0.088	0.194*	0.158***	0.17***	0.16***	0.131***
	(0)	(0.065)	(0.093)	(0.097)	(0.035)	(0.038)	(0.036)	(0.039)
mean n $servant_1$		0.026	-0.044	-0.222·	-0.146***	-0.123***	0.04	0.128*
	(0)	(0.087)	(0.1)	(0.131)	(0.034)	(0.036)	(0.05)	(0.052)

Continued on next page

vars	Homo, Naive			Homo, PML/FP	Hete, Naive		Hete, PML/FP	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
mean n servant$_2$		0.322***	0.235*	0.18	0.135***	0.085**	0.072$^{\cdot}$	0.082*
	(0)	(0.085)	(0.101)	(0.129)	(0.032)	(0.033)	(0.038)	(0.039)
mean n servant$_3$		0.229*	0.146	0.05	-0.083*	-0.138***	-0.218***	-0.209***
	(0)	(0.09)	(0.106)	(0.141)	(0.036)	(0.038)	(0.044)	(0.048)
mean n servant$_4$		0.133	0.046	-0.108	-0.196***	-0.238***	-0.166***	-0.117**
	(0)	(0.081)	(0.095)	(0.133)	(0.031)	(0.033)	(0.041)	(0.041)
mean stayer$_1$		-0.161	-0.652	-1.004$^{\cdot}$	-0.872***	-0.92***	-1.753***	-1.649***
	(0)	(0.214)	(0.502)	(0.536)	(0.232)	(0.258)	(0.254)	(0.274)
mean stayer$_2$		0.207	1.027*	1.199**	0.691***	0.557**	0.644***	0.287
	(0)	(0.19)	(0.439)	(0.453)	(0.192)	(0.213)	(0.193)	(0.216)
mean stayer$_3$		0.328$^{\cdot}$	0.595	0.415	0.419*	0.395*	-0.04	-0.017
	(0)	(0.177)	(0.419)	(0.484)	(0.182)	(0.201)	(0.217)	(0.222)
mean stayer$_4$		0.09	0.285	0.452	0.408*	0.435*	0.292$^{\cdot}$	0.134
	(0)	(0.16)	(0.378)	(0.405)	(0.165)	(0.183)	(0.176)	(0.191)
log-like	-156320	-156070	-155760	-156820	-152620	-152010	-156240	-154100
obs	165114							
x_i	yes	yes	yes	yes	yes	yes	yes	yes
z_p	no	yes	yes	yes	yes	yes	yes	yes
$\tau_{b,y}$	no	no	yes	yes	yes	yes	yes	yes
$u_{p,y}$	no	no	no	no	no	yes	no	yes

Notes: Standard errors in parentheses. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$



(a) Homogeneous x_i, z_p, t_b

(b) Heterogeneous x_i, z_p, t_b

(c) Heterogeneous x_i, z_p, t_b, t_p

Figure 2.15: Convergence of J_y^T for homogeneous and heterogeneous

2.6.5 Marginal effects

Starting from equations 2.2 and 2.3 if we consider the change in the endogenous effect we have

$$\frac{d\mathbb{P}_{y,i}}{d\mathbf{w}_i\mathbf{s}_{y^0}} = \begin{cases} J_y\mathbb{P}_{y,i}(1-\mathbb{P}_{y,i}) & \text{if } y^0 = y, \\ -J_{y^0}\mathbb{P}_{y^0,i}\mathbb{P}_{y,i} & \text{if } y^0 \neq y. \end{cases}$$

However, for the contextual individual characteristics we know a change in any of the covariates will have a direct effect on 2.2 but also an equilibrium effect through 2.3. Therefore,

$$\frac{d\mathbb{P}_{y,i}}{dx_{k,i}} = \underbrace{\frac{\partial\mathbb{P}_{y,i}}{\partial x_{k,i}}}_{\text{direct effect}} + \underbrace{\frac{\partial\mathbb{P}_{y,i}}{\partial\mathbf{w}_i\mathbf{s}_y}\frac{\partial\mathbf{w}_i\mathbf{s}_y}{\partial x_{k,i}}}_{\text{effect on beliefs of } j\in nei_i \text{ taking } y} + \underbrace{\sum_{y^0\neq y}\frac{\partial\mathbb{P}_{y,i}}{\partial\mathbf{w}_i\mathbf{s}_{y^0}}\frac{\partial\mathbf{w}_i\mathbf{s}_{y^0}}{\partial x_{k,i}}}_{\text{effect on beliefs of } j\in nei_i \text{ taking any other } y^0\neq y}$$

Doing the calculation we get

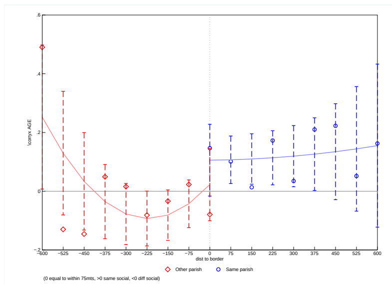
$$\begin{aligned}\tfrac{d\mathbb{P}_{y,i}}{dx_{k,i}} = \; & c_{k,y}\mathbb{P}_{y,i}\left(1-\textstyle\sum_{y'\in\Omega}\tfrac{c_{k,y'}}{c_{k,y}}\mathbb{P}_{y',i}\right)+ \\ & J_y\mathbb{P}_{y,i}(1-\mathbb{P}_{y,i})\left[\textstyle\sum_{j\in nei_i} w_{ij}\left(w_{ji}d_{k,y}\mathbb{P}_{y,j}\left(1-\textstyle\sum_{y'\in\Omega}\tfrac{d_{k,y'}}{d_{k,y}}\mathbb{P}_{y',j}\right)\right)\right]- \\ & \textstyle\sum_{y^0\neq y} J_{y^0}\mathbb{P}_{y,i}(\mathbb{P}_{y^0,i})\left[\textstyle\sum_{j\in nei_i} w_{ij}\left(w_{ji}d_{k,y^0}\mathbb{P}_{y^0,j}\left(1-\textstyle\sum_{y'\in\Omega}\tfrac{d_{k,y'}}{d_{k,y^0}}\mathbb{P}_{y',j}\right)\right)\right].\end{aligned}$$

Table 2.15: Average Marginal Effects, endogenous and exogenous variables

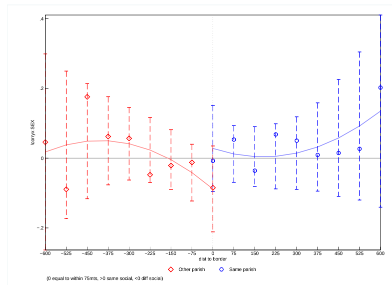
vars		**PML/FP**					
		unemployed	professional	domestic	commercial	industrial	
		(1)	(2)	(3)	(4)	(5)	(6)
Endogenous							
Estimated	$w_i \hat{s}_{y,i}$	5.37E-03	-2.41E-03	-1.83E-03	-2.46E-02	8.04E-02	
	sd $(w_i \hat{s}_{y,i})$	[0.039]	[0.048]	[0.049]	[0.054]	[0.106]	
Exogenous							sd var
	direct	-3.48E-03	1.72E-04	-2.11E-03	2.55E-03	3.71E-03	
Age	ind y	-2.16E-06	2.65E-06	-2.05E-06	-3.04E-05	1.06E-04	[10.688]
	ind $y^0 \neq y$	-1.09E-05	-7.30E-06	-1.54E-05	-6.70E-05	2.63E-05	
	direct	1.57E-03	1.41E-03	2.19E-03	3.93E-03	1.36E-02	
n child	ind y	1.52E-05	-9.48E-06	3.87E-06	2.24E-05	-4.12E-05	[1.992]
	ind $y^0 \neq y$	2.76E-06	-1.33E-05	4.17E-06	3.83E-05	-2.26E-05	
	direct	1.45E-03	1.68E-03	2.71E-03	6.50E-03	2.09E-02	
n servant	ind y	-1.39E-05	6.28E-06	6.01E-06	7.91E-06	-8.94E-05	[0.876]
	ind $y^0 \neq y$	8.22E-06	2.25E-05	2.07E-05	3.55E-05	-3.75E-06	

Notes: Standard deviation in brackets.

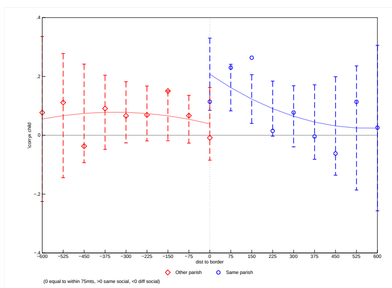
2.6.6 Discontinuities in correlations



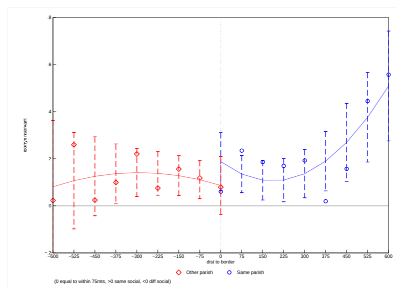
(a) age



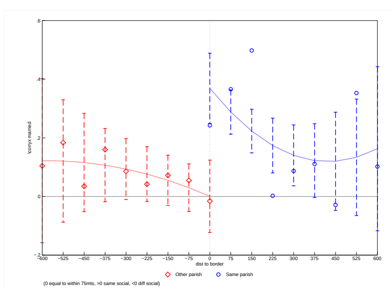
(b) male



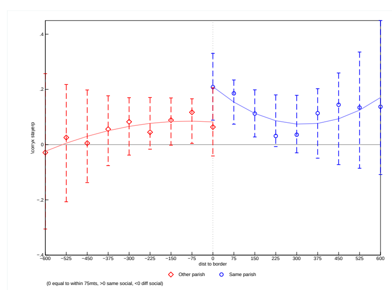
(c) children in HH



(d) number of servants

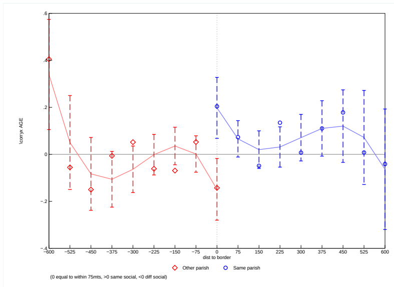


(e) married

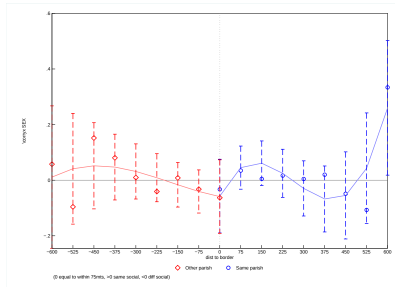


(f) parish birth=residence

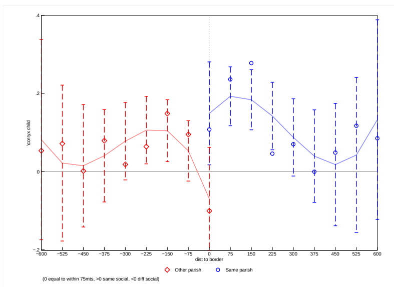
Figure 2.16: Correlation in endogenous variables for $h = 40$ mts varying neighbours at 75 mts bins, polynomial degree 2



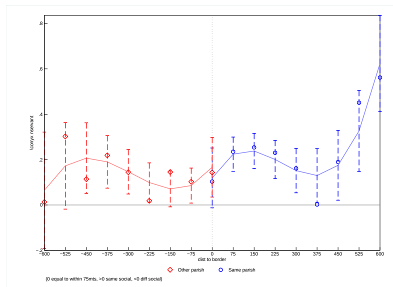
(a) age



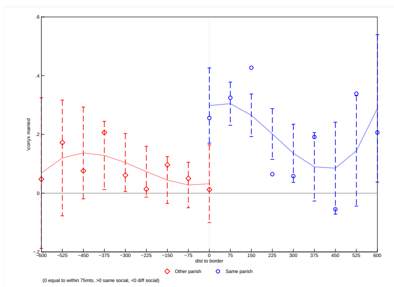
(b) male



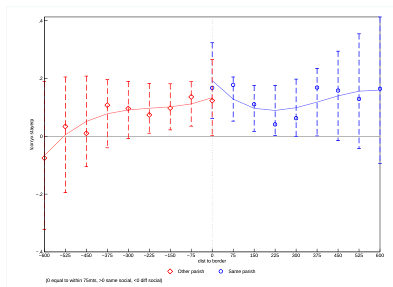
(c) children in HH



(d) number of servants



(e) married



(f) parish birth=residence

Figure 2.17: Correlation in exogenous variables for $h = 50$ mts varying neighbours at 75 mts bins, polynomial degree3

2.6.7 Administrative Areas

Table 2.16: Table of Administrative Areas

Present Borough (1965)	Metropolitan Borough (1900)	Board of Works (1855)	Civil Parishes	Ancient Parish
Camden	St Pancras	Vestry	St Pancras	St Pancras
	Hampstead	Vestry	Hampstead (St John)	Hampstead (St John)
	Holborn		St Andrew above the Bars (Holborn)	St Andrew (Holborn)
			St George the Martyr	
		Holborn District	St Andrew Holborn above the Bars with St George the Martyr	
		St Giles District	St Giles in the Fields & St George Bloomsbury (1774)	
			Staple Inn	Staple Inn Extra Parochial Place
			Furnivals Inn	Furnivals Inn Extra Parochial Place
			Grays Inn	Grays Inn Extra Parochial Place
		Holborn District	Liberty of Saffron Hill, Hatton Garden, Ely Rents & Ely Place	Liberty of Saffron Hill, Hatton Garden, Ely Rents & Ely Place & Liberty of Saffron Hill, Hatton Garden, Ely Rents & Ely place. Saffron Hill is within St Andrew Holborn
			Lincolns Inn	Lincolns Inn Extra Parochial Place
Greenwich	Greenwich	Greenwich District	Greenwich (St Alfege) Deptford St Nicholas Greenwich Deptford (St Paul)	
		Lee District (orig Plumstead)	Kidbrooke	Kidbrooke Ancient parish being regarded as liberty following loss of church and re-established in 1866
			Charlton (next Woolwich)	Charlton
	Woolwich (south of River)	Lee District (orig Plumstead)	Eltham (St John the Baptist)	Eltham
		Vestry (orig Plumstead District)	Plumstead (St Margaret)	Plumstead (originally including chapelry of East Wickham)
		Vestry	Woolwich (St Mary)	
Hackney	Hackney	Hackney Board	Hackney (St John)	Hackney
	Stoke Newington	Hackney Board	Stoke Newington (St Mary)	Stoke Newington
			Part of South Hornsey forming detached areas in Stoke Newington (parish and UD created 1896 and transferred to London in 1900)	Hornsey

Continued on next page...

Present Borough	Metropolitan Borough	Board of Works	Civil Parishes	Ancient Parish
	Shoreditch	Vestry	St Leonard (Shoreditch)	St Leonard (by 1558)
		Whitechapel District	Liberty of Norton Folgate.	
Hammersmith & Fulham	Hammersmith	Vestry	Hammersmith (St Paul)	Fulham
	Fulham	Vestry	Fulham (All Saints)	Fulham
Islington	Finsbury	Vestry	Clerkenwell	St James Clerkenwell
				St John
		Vestry	St Luke	St Giles Without Cripplegate
		Holborn District Board	Glasshouse Yard (Liberty)	St Botolph Without Aldersgate
		Holborn District Board	St Sepulchre	St Sepulchre
			Charterhouse	Charterhouse
	Islington	Vestry	St Mary Islington	St Mary Islington
Kensington & Chelsea	Kensington	Vestry	St Mary Abbots, Kensington	St Mary Abbots, Kensington
	Chelsea	Vestry	St Luke Chelsea	St Luke Chelsea
Lambeth	Lambeth	Vestry	Lambeth (St Mary)	Lambeth
	Wandsworth			
Lewisham	Deptford		Deptford (St Pauls)	Deptford
	Lewisham	Plumstead	Lee (St Margaret)	Lee
		Vestry	Lewisham (St Mary)	Lewisham
			Part of Camberwell on western slopes of Forest Hill	
Southwark	Southwark	St Saviours District Board of Works	Christchurch (Southwark)	Created parish in 1670, was originally a liberty (Paris Garden)
		St Saviours District Board of Works	St Saviour (Southwark)	Created in 1541 from the ancient parishes of St Margaret and St Mary Magdalen which were combined
		Vestry	St Mary Newington	St Mary Newington
		Vestry	St George the Martyr	St George the Martyr
	Camberwell	Vestry	St Giles Camberwell	St Giles Camberwell
	Bermondsey	Vestry	St Mary Magdalen, Bermondsey	St Mary Magdalen, Bermondsey
		Vestry voting with St Olave District	St Mary, Rotherhithe	St Mary, Rotherhithe
		St Olave District	St John Horsleydown (Southwark)	St Olave, Southwark
		St Olave District	St Olave & St Thomas (Southwark)	St Olave (Southwark)
				St Thomas (Southwark); created form area of St Olave (above) in c.1550 from area comprising Archbishop of Canterbury's hospital
Tower Hamlets	Bethnal Green	Vestry	Bethnal Green (St Matthew)	Stepney
	Poplar	Poplar District	Bow, formed 1719 from Stepney	Stepney

Continued on next page...

Present Borough	Metropolitan Borough	Board of Works	Civil Parishes	Ancient Parish
	Poplar	Poplar District	Bromley	Bromley
			Poplar (All Saints), formed 1817 from Stepney, though had been chapelry from 1654	Stepney
	Stepney	Limehouse District	Limehouse (St Anne), formed 1725 from Stepney	Stepney
		Whitechapel District	Mile End New Town, formed 1866 from Stepney	Stepney
			Mile End Old Town, formed 1866 from Stepney	Stepney
		Whitechapel District	Norton Folgate, formed 1858	Prior to 1858 was liberty and extra parochial area
		Whitechapel Distrct	Old Artillery Ground, formed 1866	Prior to 1866 was liberty
			Old Tower Without, formed 1858 and abolished 1895 (to St Botolph without)	Previously extra parochial place
		Limehouse District	Ratcliffe, formed 1866 from part of Stepney and part of Limehouse	Stepney
		Whitechapel District	St Botolph without Aldgate (being that part of St Botolph that lays outside City of London). In 1895 included Old Tower Without.	St Botolph
		Whitechapel District	St Katherine, transferred to St Botolph Without in 1895	St Katherine
			St George in the East, formed 1729 from Stepney	Stepney
		Limehouse District	Shadwell, formed 1670 from Stepney	Stepney
		Whitechapel District	Spitalfields (Christ Church), formed 1729 from Stepney	Stepney
			Stepney (St Dunstan)	Stepney
		Whitechapel District	Tower of London, created parish in 1858	Prior to 1858 was liberty and extra parochial area
		Limehouse District	Wapping, formed 1729 from part of Stepney	Stepney
		Whitechapel District (incl Holy Trinity Minories, Pr St Katherine)	Whitechapel (St Mary), formed in early 17th century from part of Stepney	Stepney
		Whitechapel District	Holy Trinity Minories, Transferred to Whitechapel in 1895	Holy Trinity Minories

Continued on next page...

Present Borough	Metropolitan Borough	Board of Works	Civil Parishes	Ancient Parish
Wandsworth	Battersea	Wandworth Board	Battersea (St Mary)	Battersea
	Wandsworth (western part)	Wandsworth Board	Clapham (Holy Trinity)	Clapham
			Putney (St Mary)	Originally chapelry of Wimbledon
			Streatham (St Leonard)	Steatham
			Tooting Graveney	Tooting Gravey
			Wandsworth (All Saints)	Wandsworth
Westminster	Westminster	Vestry	St Martin in the fields	St Martin in the fields
		Vestry	St George Hanover Square 1725	
		Vestry	St James Westminster (Piccadilly) 1685	
		Strand District	St Anne Soho 1678	
		Strand District	St Paul Covent Garden 1645	
		Westminster (1855-1885 only)	St Margaret Westminster	St Margaret Westminster
		Westminster (1855-1885 only)	St John the Evangelist Westminster 1727	
		Strand District	St Clement Danes	St Clement Danes
		Strand District	St Mary le Strand	St Mary le Strand
		Strand District Board of Works	Liberty of the Rolls	Liberty of the Rolls (a Liberty, being that part of St Dunstan's in the West situated in Middlesex)
		Strand District Board of Works	Precinct of the Savoy	Precinct of the Savoy
	Paddington	Vestry	Paddington	Paddington
		Vestry	Chelsea (det part)	Chelsea (det part)
	St Marylebone	Vestry	St Marylebone	St Marylebone

Source: `http://www.jimella.nildram.co.uk/counties.htm\#bounds`

Table 2.17: Balanced Sample across merged observations in non institutional dwellings

	Not merged			Merged			$H_0 : 1 - 2 = 0$	
	N	Mean (1)	SE	N	Mean (2)	SE	t-stat	pval
All individuals								
Male	802,735	0.450	0.001	895,712	0.473	0.001	30.430	0.000
Age	802,746	32.521	0.014	895,718	32.767	0.013	13.104	0.000
Pop Age 25-34	802,746	0.279	0.000	895,718	0.276	0.000	-4.087	0.000
Pop Age 35-44	802,746	0.207	0.000	895,718	0.205	0.000	-3.282	0.001
Pop Age 45-60	802,746	0.194	0.000	895,718	0.203	0.000	14.673	0.000
Native	802,746	0.934	0.000	895,718	0.920	0.000	-32.871	0.000
Labour Force	802,202	0.665	0.001	894,880	0.694	0.000	40.229	0.000
Married	799,622	0.590	0.001	892,655	0.608	0.001	24.750	0.000
Individuals in non extreme border parishes†								
Male	150,171	0.473	0.001	491,167	0.467	0.001	-3.737	0.000
Age	150,171	32.822	0.032	491,170	32.731	0.018	-2.499	0.012
Pop Age 15-24	150,171	0.319	0.001	491,170	0.320	0.001	0.574	0.566
Pop Age 25-34	150,171	0.266	0.001	491,170	0.273	0.001	5.446	0.000
Pop Age 35-44	150,171	0.208	0.001	491,170	0.202	0.001	-4.605	0.000
Pop Age 45-60	150,171	0.207	0.001	491,170	0.205	0.001	-2.065	0.039
Native	150,171	0.904	0.001	491,170	0.901	0.000	-2.316	0.021
Labour Force	150,017	0.700	0.001	490,731	0.708	0.001	6.204	0.000
Married	149,572	0.600	0.001	489,443	0.590	0.001	-6.532	0.000
Individuals in non border parishes‡								
Male	97,259	0.489	0.002	293,892	0.488	0.001	-0.523	0.601
Age	97,259	32.835	0.040	293,894	32.761	0.023	-1.602	0.109
Pop Age 15-24	97,259	0.318	0.001	293,894	0.322	0.001	1.968	0.049
Pop Age 25-34	97,259	0.264	0.001	293,894	0.267	0.001	1.532	0.126
Pop Age 35-44	97,259	0.210	0.001	293,894	0.204	0.001	-3.895	0.000
Pop Age 45-60	97,259	0.207	0.001	293,894	0.207	0.001	-0.052	0.959
Native	97,259	0.896	0.001	293,894	0.891	0.001	-3.615	0.000
Labour Force	97,160	0.701	0.001	293,560	0.703	0.001	1.211	0.226
Married	96,866	0.615	0.002	292,865	0.616	0.001	0.863	0.388

‡ *border parishes* are: Battersea, Bow, Bromley St Leonard, Brompton, Camberwell, St Dunstan Stepney/Mile End, Mile End New Town, Poplar, St George Hanover Square, St James Clerkenwell, St Leonard Shoreditch, St Luke Chelsea, St Luke Old Street, St Margaret Westminster, St Mary Abbots Kensington, St Mary Lambeth, St Mary Paddington, St Mary Rotherhithe, St Marylebone, St Matthew Bethnal Green, St Nicholas Deptford, St Pancras, St Paul Deptford, Greenwich, St Anne Kensington, Brompton, Islington. † extreme border parishes are: *border parishes* minus St George Hanover Square, St Luke Old Street, St Dunstan Stepney/Mile End, Mile End New Town, St Marylebone.

Chapter 3

Reference points formation and investment decisions

3.1 Introduction

The ideas delivered under the prospect theory framework by Kahneman & Tversky (1979) have gain increasing relevance in the economic literature. Many experimental results have showed that a loss, with respect to a reference point, is more painful to individuals than a commensurable gain. Also there is evidence that marginal sensitivity of further deviations is decreasing both below and above the reference point. This translates into risk–loving behaviour in losses while being risk averse in gains. However, *what* determines the reference point remains an open question.

Early suggestions on what a decision maker reference point is were mainly focused on *backward looking* introspection and *unique* reference points (standard Prospect Theory, PT Kahneman & Tversky (1979)). In that sense, any form of current status–quo, or of habits formation were suggested as a determinant of the reference point. In all such cases, actual experiences feed the benchmark against which individuals weight the outcome from their decisions. More recently, some authors have suggested that expectations that a person held when start focusing on the decision and shortly before the outcome occurs (Expectation Based Prospect Theory, EBPT (Köszegi & Rabin

[0]We acknowledge the comments from Syngjoo Choi, Erik Eyster, Matthew Rabin, Charles Sprenger, Georg Weiszäcker, Francesco Cerigioni and Lukas Wenner. The present version benefit from audiences at the EEA-ESEM Gothenburg Meeting 2013, the BEElab Florence 2013, THEEM Konstanz 2013, UCL work in progress seminar, and ULB Jamboree Meeting 2013. Without the contribution from Ricardo Argüello, Bernardo Atuesta, Mariana Blanco, Juan–Camilo Cardenas, Diana Contreras, Jenny Espinosa, Carolina Lopera, Manuel Ramírez, Mauricio Rodríguez, Oscar Salazar, Renata Samacá and Hernando Zuleta this chapter would not have been written. Financial support from the Colombian Central Bank and Colciencias grant 1222-405-20240, CT 144-2007 is appreciated.

2006)) may affect reference point formation. In these later cases reference points are generally *multiple* and *forward* looking.

In the present chapter we recognize that recently held expectations on both: possible consumption levels (i.e. counterfactuals) and actual experienced consumption level affect individual decision making. Therefore, foregone opportunities as well as actual experiences serve as benchmarks against which individuals may compare prospective outcomes from current actions. We thus allow for multiple reference points that take as inputs lagged beliefs and lagged status quo. We apply such set up on an investment decision framework.

Specifically, in this chapter we provide a model on investment decisions with multiple reference points combining stochastically expectations on forgone outcomes together with actual outcomes. The model delivers some behavioural results to be tested experimentally against competing theories of reference points formation. We draw our motivation from the fact that it is not the same deciding how much to invest on a risky alternative from a certain available income which one has earned after a winning event, than starting with the same amount after losing the chance of starting with a higher one.

None of the previous papers on investment decisions allow for beliefs to play a role in determining the reference point, and therefore they are usually deterministic with an adjustment based on starting conditions and recent outcomes. Our approach fills this gap.

An important theoretical result we find is that willingness to invest in risky options differs non–monotonically across income levels, specifically, the low income individuals are willing to bear more risk than high–initial income individuals when stochastic reference points are introduced. In this chapter we also provide an experiment to test some of the implied behavioural results on a sample of informal small entrepreneur.

In the following section we provide a short overview of related work and then introduce the experimental set–up we use to test the theoretical framework presented in section 3.4. The model incorporates stochastic reference–dependent preferences into an investment decision problem. In section 3.5 we present the econometric results from the experiment and test the theoretical implications of the proposed model against sensible competing theories on reference point formation.

3.2 Related Literature

The EBPT perspective seems to be consistent with some results found in the experimental literature when individual's status quo is not held constant. Abeler, Falk, Goette & Huffman (2011) experimental set-up suggest that an expectation-based approach is sensible in determining effort provision. Also, Ericson & Fuster (2011) provide evidence that expectation–based reference points affect exchange behaviour. Köszegi & Rabin (2007) also show that endogenous reference points driven by beliefs can affect risk attitudes, they state that a person is less risk averse in eliminating a risk she expected to face than in taking on the same risk if it was unexpected. Sprenger (2010) tests specifically whether reference points can be determined by expectations. He is interested in contrasting competing theories of reference point determination: EBPT, Disappointment Aversion (DA, Loomes & Sugden (1986)) and EUT. His experimental test suggest the existence of an endowment for risk which is uniquely predicted by EBPT.

Evidence on path–dependence risk attitudes is given by Post, Van den Assem, Baltussen & Thaler (2008). Their findings suggest that individual decisions cannot be fully reconciled with Expected Utility Theory (EUT) preferences. In particular they find that risk aversion is affected by previous outcomes experienced during the game. Those experiencing large paper losses earlier exhibit a "break–even" effect reducing their risk aversion in later rounds, which suggests that their reference point adapts somewhat slowly to previous losses; those experiencing paper gains, as well, reduce their risk aversion in later rounds.

There have been several papers that built upon reference points framework to investigate individual decisions. Bowman, Minehart & Rabin (1999) propose an axiomatic approach to study how the inclusion of references points and loss aversion affects decisions over savings and consumption in a two-period framework with habit formation. They assume no borrowing constraints, no discount rate and savings earn no interest; the uncertainty in their model is on income earned in the second period. They find empirical and theoretical evidence that, given an expected income above (below) the reference point, an increase in uncertainty leads the consumer to increase (decrease) savings. Siegmann (2002) takes an especial case of the value function without habit formation nor discount rate, but with uncertainty over interest rates. She finds that the saving function is non-decreasing in income and differs across poor and rich.

Berkelaar et al. (2004) provide a model on investment decisions and portfolio allocation with reference dependent preferences. They assume a complete market framework and find that loss averse investor will follow a partial portfolio insurance strategy, giving up wealth in good states in order to keep wealth above the reference point at the planning horizon. This strategy reduces the initial portfolio weight of stocks. Gomes (2005) studies instead the implication for trading volumes with reference dependent preferences, he finds that demand function for risky assets is discontinuous and non–

monotonic: once wealth reaches to a certain threshold individuals will insure themselves against possible losses with respect to a deterministic reference point subject to habit formation, this happens mainly because their utility is defined only on the gain–loss domain.

On the other hand, Yogo (2008) investigates the effect of reference-dependent preferences on asset prices which allows explaining the high equity premium found in the data and its countercyclical behaviour. This is done by assuming that reference points are based on slow changing habits.

3.3 Experimental Design

The experiment was implemented during October and November 2008. Overall we ran 13 sessions within seven days which amounts to two sessions per day, one from 7 am to 9 am, and the other one from 5pm to 7 pm and an average of 15 subjects per session. Our subject sample is composed by informal small entrepreneurs and amounts to 216 members of the Institute for Social Economy, Colombia (IPES in spanish).These entrepreneurs are mostly artisans and street vendors who are very vulnerable to shocks and are constantly taking decisions under uncertainty that affect their future income.

From our sample of entrepreneurs, 64.8% are female within 20 and 50 years. Nearly 25.9% of our whole sample have less than high school degree, 30% have completed secondary education, 25.9% have technical education and less than 6% have completed a higher education degree. The monthly income of almost 50% of this population is less than 5 minimum wages (i.e. approximately USD 200 by 2008) and 80% do not save for retirement.

Experimental sessions were implemented in the IPES facilities so subjects did not need to commute to the university lab and therefore were in a more familiar working environment. These allowed subjects to approach experiment's decisions more naturally than had it been in an experimental laboratory within university premises. All the instructions where read aloud with the help of a video–beam to guarantee each participant got the same understanding of the instruction and experimental procedures.

Given subjects did not have much familiarity with computers and decisions at stake were cumbersome, the implementation of each session was in charge of 15 undergraduate students hired and trained specifically for the project, we called them Monitors. The Monitors knew the purpose of the whole experiment and where given detailed instructions on the wording they should use across the experiment in order not to influence any behaviour.[1] One Monitor was allocated to each subject, they were in

[1] See appendix 3.7.3 for the instructions handed in to the Monitors.

charge of entering the decisions from the subjects into the computer and implementing any randomized draw so participants were certain of it being transparent.

In addition, as the experiment was long and aimed at vulnerable population we gave them, at the beginning, some refreshments so they could endure the whole procedure. This prevented them from rushing through the final stage of the procedure. We provided them with a folder, labelled with the university logo, which contained the preliminary instructions, a socio–demographic survey and the instructions from the first stage. As soon as the first stage was completed, they were told they would proceed to the second stage without knowing the pay–off from the previous stage and that they would receive the total experimental pay–off at the end of the session.

3.3.1 Procedure

In the overall experiment there were two stages. In the first stage we rationalize the individual choices over a paired sequence of price–list lotteries following the experimental design of Tanaka, Camerer & Nguyen (2010) to determine parametrically each individual Prospect Theory parameters: overweighting of small probabilities (which we denote α), loss aversion (λ) and risk aversion (γ). In the appendix we include a detailed description of it.

It is important to note that none of the decisions taken in the first stage affected the decisions of the next stage. Experimental pay–offs were paid only after finishing both stages. Subjects were clearly informed that none of the decisions made at that point affected outcomes from the next stage and that answers were not considered either correct or incorrect. We also collected a standard socio economic characterization survey.

At the beginning of the second stage individuals completed a socio economic survey, once they have done this we gave them an income, depending on the treatment they were included in (more details below). The subjects were told that the assigned income was of their own already because they had reached to this point of the experiment. This is important for our implementation because we managed to make subjects feel that the assigned income belonged to them already. Given an initial income, subjects are face with investment alternatives and must decide what to invest on a risky alternative and a risk–free one. All the instructions are to be found in the appendix 3.7.3.

Treatments

We split the population into four treatments that were characterised by an *Income Variation (IV)* and a *Reference point Variation (IV)*. The *RV* splits the population in half according to how the initial income would be generated while; the IV determines how high that initial income would be.

In particular, the initial income level (ω) could take two values: High $\overline{w} = COP40,000$ (i.e. USD 20)[2] or low $\underline{w} = COP20,000$ (i.e. USD 10).

$$\omega = \begin{cases} \underline{w} & \text{Low initial income, } \simeq USD\ 10, \\ \overline{w} & \text{High initial income, } \simeq USD\ 20. \end{cases}$$

The RV vary how that initial income would be generated. We split the sample into a first group of individuals (L_0) that started the experiment with a predetermined income. Importantly, they were not informed about the possibility of getting a different income level than the one assigned to them. A second group of individuals (L_1) were informed there were two possible income levels one of which would be assigned to them depending on a coin toss. They were then given a coin for them to choose which side they wanted to associate with the high (low) level and to toss it afterwards to determine their own income.

$$L = \begin{cases} L_0 & \text{predetermined } \omega, \\ L_1 & \text{coin toss lottery over feasible } \omega's. \end{cases}$$

For all treatments, after the income level was assigned, we show the subjects the physical note and remind them that it was already theirs, then leave it on the table in front of them.

We were very careful on making the allocation and randomization of treatments very transparent to avoid confounding effects on subjects' decisions. The Table 3.1 summarize the four treatments together with the distribution of observations in our sample. Subjects starting with a predetermined income of $USD10$ represent 35.65%, those that started with a high income level $USD20$ are 33.80% of the overall sample. The rest of the population face the coin toss to determine their initial income level and they were evenly distributed between high and low income levels.

Once individuals receive the initial income they will then take a series of investment decisions across several rounds t and states of the wolrd s. In particular, at moment t individuals decide how much, of an available income w^t, to invest in a single risky alternative if state of the world is s. We give more details in the following section

[2] Equivalently to 10% of the 2008's Monthly Minimum Wage in Colombia.

Table 3.1: Experimental treatments (L, ω)

$L \backslash \omega$	$\underline{\mathbf{w}}$	$\overline{\mathbf{w}}$	freq.
$\mathbf{L_0}$	Fixed Low income	Fixed High income	150
$\mathbf{L_1}$	Random Low income	Random High income	66
TOTAL	50.0%	50.0%	216*

* Total decisions are 5400. Each individual takes 5 investment decisions across 5 rounds .

Decisions

After individuals receive a starting income $\omega = w^0$ (round $t = 0$) they were asked to take two set of decisions across $T = 5$ different rounds which are linked to each other in the sense that final income in round t is the starting income at $t+1$. The number of rounds are not known, ex–ante, to the individuals. Instead of trying to capture some effects of infinite horizon dynamic problem we gave this instruction in order to remove any specific concern from what to expect from future rounds and therefore, round by round, the only relevant ex–ante difference is the available income at the beginning of each round.

First decision: Investment $\theta_{t,s}$

Given initial income, they have to decide what part of it (i.e. $\theta_{t,s} \in [0,1]$) they are willing to invest in a Risky Alternative when the rate of return depend on equally likely random states of the world $s \in \{1, \cdots, 5\}$. Each state is associated to a rate of return $\overline{r}_1 = 0,05; \overline{r}_2 = 0,10; \overline{r}_3 = 0,15; \overline{r}_4 = 0,20$ or $\overline{r}_5 = 0,30$ if the outcome of the Alternative is successful, and $\underline{r}_s = -0.90$ for all s if unsuccessful. They were told that each rate of return $r_s, s \in \{1, \cdots, 5\}$ was supposed to be considered as totally independent from each other because only one state of the world would be chosen at the end of a given round. So that, within each round, they had the same total income to be invested in each state.

Any amount not invested is left for next decision round, which is equivalent to allocate that amount on a Safe Investment with a rate of return equal to 1. Individuals were reminded that the assigned income was their own already, therefore a perfectly viable decisions was to keep all of it in the Safe Investment round after round.

Second decision: Insurance $D_{t,s}$

After they have revealed The second decision individuals make at each round t is whether they want to keep the Option for the Risky Investment (that we called *Option A*) or change it for a Fair Insurance at each of the states of nature $s = \{1, \cdots, 5\}$ that we labeled *Option B*. We denote this decision as $D_{t,s} \in \{A, B\}$

Round by round

As the Risky Investment has two possible outcomes: success or failure, we allow the likelihood of an event to change across rounds. Success probability ρ^t is set to be decreasing across rounds: $\rho^1 = 0.1, \rho^2 = 0.3, \rho^3 = 0.5, \rho^4 = 0.7, \rho^5 = 0.9$, a pattern not ex–ante known by subjects. This, instead of bringing uncertainty or ambiguity about subjects' expectation regarding success probabilities in the future, help us to be sure that individuals were revealing their willingness to invest at each probability of success rather than thinking strategically on whether to wait for a round until the deterministic rate of return is maximum.[3]

In each round t, after individuals had revealed their two decisions in each state s, we throw a five–sides die to determine the state of nature to be taken into account. If individual's decision was to keep *Option A* the subject was ask to randomly draw a ball from a bag in which 10 balls, a combination of green and red colored balls, had been included already representing the corresponding probability of success of the business, where the green color was associated to the business being successful. Otherwise, if *Option B* had been chosen we updated automatically individual's income.

Notice that at the end of each round t, an individual have a final income, w^{t+1} computed by taking into account the random state of the world, the amount invested in the risky alternative, the decision between staying with the investment or accepting a fair insurance and the amount left in the safe option.

3.3.2 Summary

Figure 3.1 summarizes the experimental procedure

At each round t an individual starts her decisions with an income w^t, in each round (t) she is informed about the probability of success (ρ^t) associated with the risky alternative. First, individuals must decide what part of their income to invest in the risky alternative given state s ($\theta_{t,s}$) and what part they wish to invest in the safe alternative ($1 - \theta_{t,s}$), where s represents a particular state of the nature that draws a particular rate of return $\overline{r}_s$ if successful. If the risky investment is unsuccessful the investor would

[3]However, a cleaner treatment would have been to allow for the rate of return to be randomly chosen, with given support, round by round. In the analysis section we asses whether some of the subjects were able to identify a deterministic pattern on the probability of success across consecutive rounds by studying their decisions round by round by interacting indicator treatment variables with round dummies, we did not find any evidence of a statistical significant effect,

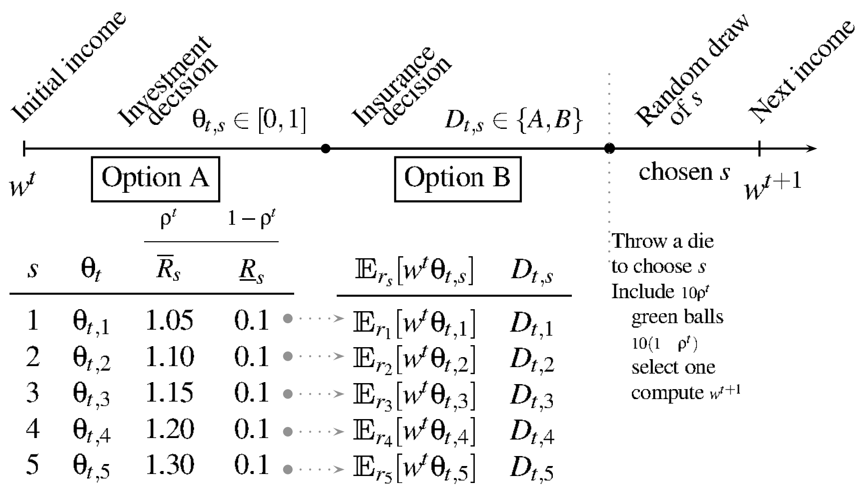


Figure 3.1: Decision timing for a given round t, with probability of success ρ^t

lose 90% of the invested amount, no matter the state of nature. Individuals know each state s is independent of each other, therefore at a given round t the available income w^t is the same for each investment decision. Then individuals are given the option to decide ($D_{t,s}$) whether they want to go through with the stated investment (*Option A*) or change it for a fair insurance given her invested amount (*Option B*). After these decisions, we randomly choose the relevant state s and compute final income w^{t+1} for that round according to the prerecorded decision from the individual. This final income will be the initial income at the following round. At the end of all rounds we calculate accumulated payments. All individuals face a practice round before the actual experiment starts.

This experimental implementation would allow us to contrast the theoretical predictions included in the next section and their derivations in the appendix. Our approach on the reference point formation assumes that it is determined by the expectations held once individual focus on the decision at stake. In this case, an individual that got an income level without knowing the existence of an alternative one, will associate the very income level with her reference point. On the contrary, an individual who knew about the existence of two income levels and the probability of getting them, will form her reference point as either an intermediate point, consistent with the equal probability of getting them, or will weight outcomes with mixed feelings against the two levels.

3.4 Theoretical framework

Our baseline framework is built on the prospect theory elaborated by Kahneman & Tversky (1979). The implemented treatments give us some power of discriminating several competing theories of decision making under uncertainty: EUT, non–stochastic reference–dependent model (i.e. standard Prospect Theory, PT), Prospect Theory with Stochastic reference points (SPT) and Sampling–Based Prospect Theory (SBPT). SPT follows closely the appraisal of Köszegi & Rabin (2006) for surprise decisions where the stochastic reference points are given exogenously.

Our modelling framework assumes that an agent has an initial income ω and must decide what proportion θ to invest in a risky asset and how much to invest on a safe asset with return R_f. The outcome of the investment is either Success (S) or Failure (F). The rate of return if successful is $\overline{R}$ with probability ρ and unsuccessful with complimentary probability. The problem here considered is static. This is reasonably given that in the experimental treatments subjects did not know the number of decision round they were going to face and therefore every round was regarded as the last one.

The final wealth $y = (1-\theta)\omega R_f + \theta\omega R$, which implicitly depends on $\left(\omega, \overline{R}, \underline{R}, R_f, \theta\right)$[4], is consumed c. The utility is captured by $u\,(c|b) \equiv m\,(c) + \mu\,(c|b)$ where $m\,(c)$ is the usual outcome-based utility, which is strictly increasing and concave. b is her reference point or benchmark. The more general form for the second term is one in which both consumption and reference point are translated into values attached to the goods, in this sense $\mu\,(c|b) \equiv v\,(m\,(c) - m\,(b))$.

For the purpose of the current exposition we assume the simple piece-wise version for the gain–loss utility $\mu\,(c \mid b)$, that is

$$\mu\,(c \mid b) \;=\; \begin{cases} m(c) - m(b) \text{ if } c \geq b \\ -\lambda\,(m(b) - m(c)) \text{ if } c < b. \end{cases} \tag{3.1}$$

Our gain–loss utility is well behaved as is standard in the literature. It satisfies desirability of the consumption no matter which is the reference point and the marginal utility of a loss is strictly greater to the marginal utility of a commensurable gain. The parameter λ can be interpreted as the weight attached to a gain-loss utility. If $\lambda = 1$ every gain has exactly the same decisional weight attached than a commensurable loss, the higher is λ the more pain or distaste an individual suffers from a loss with respect to a comparable gain.

From now on our interpretation when referring to b is that it represents a thresh-

[4] Notice then that y is a random variable with conditional distribution function $F(\cdot \mid \theta)$ which with probability ρ gives $(1-\theta)\omega R_f + \theta\omega\overline{R}$ and with complimentary probability $(1-\theta)\omega R_f + \theta\omega\underline{R}$.

old against which an individual compares her consumption associated to a recently held belief on the income she could have attain. Therefore, we introduce exogenous stochastic reference points in our model that keep close relation to the experimental setup. Later on we evaluate the implication of different income levels in the optimal investment decision.

For easiness of exposition let us denote with $\theta_M^{\omega,L}$ the predicted invested amount on the risky asset for those individuals under Income Variation $\omega \in \{\underline{w}, \overline{w}\}$, Reference Variation $L \in \{L_0, L_1\}$ and approach $M \in \{EUT, PT, SBPT, SPT\}$ regarding formation of b.[5] That is

$$\theta_M^{\omega,L} \in \arg \max_{\{\theta:\theta\in[0,1],c\leq y\}} U_M(c \mid \omega, L). \tag{3.2}$$

3.4.1 Expected Utility Theory.

Under EUT we assume that individual preferences are based solely in the out-come–based utility $m(c)$, which amounts to assuming $\lambda = 0$, and therefore individuals decide what to invest based on the following function, were we have incorporated the randomness of consumption $c = y$ captured by $F(y \mid \theta)$

$$\theta_{EUT}^{\omega,L} \in \arg\max_{\theta} \int m(y)\, dF(y \mid \theta). \tag{3.3}$$

it is ready noticeable that experimental variation L won't have any effect on decisions under this framework, and variation ω will have an effect driven solely by the assumption imposed by the function $m(\cdot)$.

Specifically, the first order condition is characterized by

$$\rho(\overline{R}-R_f)\omega m'(\omega(R_f+(\overline{R}-R_f)\theta))-(1-\rho)(R_f-\underline{R})\omega \underline{r} m'(\omega(R_f-(R_f-\underline{R})\theta)) \begin{cases} \leq 0 & \text{if } \theta & <1 \\ \geq 0 & & >0 \end{cases} \tag{3.4}$$

and a necessary condition for there to be an interior optimum, given $m'' \leq 0$, is that $\frac{\rho(\overline{R}-R_f)}{(1-\rho)(R_f-\underline{R})} \geq 1$.

Denote $\theta_{EUT}^{\omega,L}$ the optimal solution to the previous FOC whenever initial income is $\omega \in \{\underline{w}, \overline{w}\}$ and it was determined by lottery $L \in \{L_0, L_1\}$. We know that $\theta_{EUT}^{\omega,L} = 0$ if $\frac{(\overline{R}-\underline{R})}{(R_f-\underline{R})} < \frac{1}{\rho}$ and $\theta_{EUT}^{it,rt}$ implied by equality in 3.4 otherwise.

Notice first that

Proposition 3.4.1. *For any $m(x)$ such that $m'(x) > 0$ and $m''(x) \leq 0$ then $\theta_{EUT}^{\omega,L_0} = \theta_{EUT}^{\omega,L_1}$ for all ω.*

[5] It implicitly depends on $(\overline{R},\ \underline{R}, R_f, \rho)$

The previous claim follows easily from the fact that treatments do not impose any difference between any of the investment decision lotteries. By the time the individuals are facing investment decisions the coin toss of the initial income has already passed, and therefore any of the two reference variations L are equivalent decision problems.

Additionally we have the following proposition, where $r'_R(x \mid m)$ is the relative risk aversion associated to utility m and income x , we provide the proof of this and subsequent claims in appendix 3.7.1

Proposition 3.4.2. *If $r'_R(x \mid m) \leq 0$, $m'(x) > 0$ and $m''(x) \leq 0$ then $\theta_{EUT}^{\overline{w},L_0} \geq \theta_{EUT}^{\underline{w},L_0}$.*

which is the well known result of increases on invested amounts for larger income levels when there is decreasing relative risk aversion

Proposition 3.4.3. *If $r'_R(x \mid m) \leq 0$, $m'(x) > 0$ and $m''(x) \leq 0$ then $\frac{\partial \theta_{EUT}^{\omega,L}}{\partial \rho} \geq 0$ and $\frac{\partial \theta_{EUT}^{\omega,L}}{\partial \overline{R}} \geq 0$.*

The proof of the previous result uses the implicit theorem as in proposition 3.4.2.

Summary EUT *Under EUT we should not see any difference across Reference Point variations for individuals with same income levels, no matter the values of ρ or $\underline{R}$, $\overline{R}$.*

3.4.2 Prospect Theory.

In the standard PT framework we allow for $\lambda \neq 0$ and define the reference point b as the income received, ω, multiplied by the risk free return R_f, meaning that the initial income becomes individual's status quo. The optimal decision is given by

$$\theta_{PT}^{\omega,L} \in \arg\max_{\theta} \int u\left(y \mid \omega R_f\right) dF\left(y \mid \theta\right). \tag{3.5}$$

However, as it was the case with EUT, variation in L does not change the prediction for different ω–variations. More precisely, individuals under treatment (ω, L_0) should behave as if they were under (ω, L_1), for status quo for both individuals is the same value equal to ωR_f. If there is any difference among income variation $\omega = \overline{w}$ and $\omega = \underline{w}$ it is entirely driven by $m(\cdot)$.

The first order condition to decision problem 3.5 is given by

$$\rho 2\omega\overline{r}m'(\omega(1+\theta\overline{r})) - (1-\rho)(1+\lambda)\omega\underline{r}m'(\omega(1-\theta\underline{r})) \begin{cases} \leq 0 & if & < 1 \\ \geq 0 & & > 0 \end{cases} \tag{3.6}$$

Which we know $\theta^{\omega,L_0}_{PT} = 0$ if $\frac{(\overline{R}-\underline{R})}{(R_f-\underline{R})} < \frac{1}{2\rho}(\lambda+1-(\lambda-1)\rho)$ and $\theta^{\omega,L_0}_{EUT}$ implied equality in expression 3.6 otherwise, for all ω.

Notice however that those starting with the lottery to determine their initial income (variation $L = L_1$) will solve the same decision problem 3.7

$$\begin{aligned} \underset{0\leq\theta\leq1}{Max}\,\rho\left[m(\omega(R_f+(\overline{R}-R_f)\theta))+\left(m(\omega(R_f+(\overline{R}-R_f)\theta))-m(b)\right)\right]+\\ (1-\rho)\left[m(\omega(R_f-(R_f-\underline{R})\theta))-\lambda\left(m(b)-m(\omega(R_f-(R_f-\underline{R})\theta))\right)\right] \end{aligned} \tag{3.7}$$

and therefore we have

Proposition 3.4.4. *For any $m(x)$ such that $m'(x) > 0$ and $m''(x) \leq 0$ then $\theta^{\omega,L_0}_{PT} = \theta^{\omega,L_1}_{PT}$ for all ω.*

And also the results encompassed in proposition 3.4.3 will hold. In summary, Standard Prospect Theory and EUT do not predict any difference across Reference Point Variation, and differences across Income Variation are determined by the consumption utility ($m(.)$). In particular,

Proposition 3.4.5. *If $r'_R(x \mid m) \leq 0$, $m'(x) > 0$ and $m''(x) \leq 0$ then $\theta^{\overline{w},L_0}_{PT} = \theta^{\underline{w},L_0}_{PT}$*

The proof follows the same procedure as Proposition 3.4.2 The only relevant difference between PT and EUT decision makers is that former individuals are less willing to invest than later ones with same starting incomes. This is so because of first–order risk aversion around reference point from PT decision makers.

Summary PT *(i) For PT we should not expect any difference across Reference Point variations, and differences across Income variations are determined by the outcome–based utility and therefore, follow the same direction as EUT. (ii) The only relevant difference between PT and EUT decision makers is that former individuals are less willing to invest than latter ones with same starting incomes.*

As we depict below, taking a stochastic reference point or sampling–based reference–dependent preferences approach we get results that distinguish between reference variations. In particular, no matter the starting income level, individuals who face the initial lottery are at least equally willing to invest in the risky alternative than those who did not face the lottery. Individuals generally compensate losses, from investing in the risky asset, with respect to the high reference point, with gains when weighing that decision against the low reference point.

3.4.3 Stochastic reference point

In a more general framework, where we take into account how recently held beliefs play a role in determining reference points, we have that utility is given by the average of $u\left(c|b\right)$ generated by all feasible b's under a distribution $G_L\left(\cdot\right)$. This distribution represents the beliefs an individual had about possible reference points she might have faced. Therefore

$$\theta^{\omega,L}_{SPT} \in \arg\max_{\theta} \int\int u\left(y|b\right) dG_L\left(b\right) dF\left(y \mid \theta\right). \tag{3.8}$$

Under SPT we assume that $G_L\left(b\right)$ captures the recently held beliefs about reference points. Relating to the experimental implementation we will have that $G_{L_1}\left(b\right)$ is a binary distribution, therefore $b = \overline{w}R_f$ or $b = \underline{w}R_f$ with probability $\frac{1}{2}$ given the fair–coin toss. While G_{L_0} is a degenerated distribution such that $b = \omega R_f$ with probability 1.

For an individual under treatment $(\underline{w}, L_1)$ then $\overline{b} \equiv \overline{w}R_f > \underline{w}(1+\theta\overline{r}) > \underline{b} \equiv \underline{w}R_f > \underline{w}(1-\theta\underline{r})$ and therefore, he seeks to maximize

$$\begin{aligned}
\underset{0\leq\theta\leq1}{Max}\,(1-\pi)\,&\left\{\rho\left[m(\underline{w}(R_f+(\overline{R}-R_f)\theta))-\lambda\left(m(\overline{b})-m(\underline{w}(R_f+(\overline{R}-R_f)\theta))\right)\right]+\right.\\
&(1-\rho)\left[m(\underline{w}(R_f-(R_f-\underline{R})\theta))-\lambda\left(m(\overline{b})-m(\underline{w}(R_f-(R_f-\underline{R})\theta))\right)\right]\Big\}+\\
&\pi\left\{\rho\left[m(\underline{w}(R_f+(\overline{R}-R_f)\theta))+\left(m(\underline{w}(R_f+(\overline{R}-R_f)\theta))\right)-m(\underline{b})\right]+\right.\\
&(1-\rho)\left[m(\underline{w}(R_f-(R_f-\underline{R})\theta))-\lambda\left(m(\underline{b})-m(\underline{w}(R_f-(R_f-\underline{R})\theta))\right)\right]\Big\}
\end{aligned} \tag{3.9}$$

Assuming $\pi = 1/2$, the FOC and SOC of equation 3.9 imply $\theta^{\underline{w},L_1}_{SPT} = 0$ if $\frac{\rho(\overline{R}-R_f)}{(1-\rho)(R_f-\underline{R})} < \frac{2(1+\lambda)}{3+\lambda}$, otherwise $\theta^{\underline{w},L_1}_{SPT}$ is implied by the following equation

$$\frac{3+\lambda}{2}\rho(\overline{R}-R_f)m'(\underline{w}(R_f+(\overline{R}-R_f)\theta))-(1+\lambda)(1-\rho)(R_f-\underline{R})m'(\underline{w}(R_f-(R_f-\underline{R})\theta))=0 \tag{3.10}$$

For individuals facing treatment $(\underline{w}, L_0)$ we know their decision problem is just given by equation 3.7 because $b = \underline{w}R_f$ $\theta^{\underline{w},L_0}_{SPT} = \theta^{\underline{w},L_0}_{PT}$. We can claim then that

Proposition 3.4.6. *If $r'_R(x \mid m) \leq 0$ then $\theta^{\underline{w},L_1}_{SPT} > \theta^{\underline{w},L_0}_{SPT}$ for any $\frac{(\overline{R}-\underline{R})}{(R_f-\underline{R})} > \frac{2(1+\lambda)-(\lambda-1)\rho}{(3+\lambda)\rho}$ and $\theta^{\underline{w},L_1}_{SPT} = \theta^{\underline{w},L_0}_{SPT}$ otherwise.*

Now, for subject under treatment $(\overline{w}, L_1)$ we have $\overline{w}(1+\theta\overline{r}) \geq \overline{b} \geq \overline{w}(1-\theta\underline{r}) > \underline{b}$ if $\theta \in [0, \frac{R_f}{2(R_f-\underline{R})}]$ and $\overline{w}(1+\theta\overline{r}) \geq \overline{b} > \underline{b} \geq \overline{w}(1-\theta\underline{r})$ if $\theta \in (\frac{R_f}{2(R_f-\underline{R})}, 1]$. We know a global maximum is found whenever the argument is within $[0, \frac{R_f}{2(R_f-\underline{R})}]$[6] therefore individuals under such treatment solve

[6]The proof follows the same procedure as Berkelaar et al. (2004). First we compute the conditions

$$\begin{aligned}
\underset{0\leq\theta\leq\frac{R_f}{2(R_f-\underline{R})}}{Max} \; & \pi\left\{\rho\left[m(\overline{w}(R_f+(\overline{R}-R_f)\theta))+\left(m(\overline{w}(R_f+(\overline{R}-R_f)\theta))-m(\overline{b})\right)\right]+\right.\\
& (1-\rho)\left[m(\overline{w}(R_f-(R_f-\underline{R})\theta))-\lambda\left(m(\overline{b})-m(\overline{w}(R_f-(R_f-\underline{R})\theta))\right)\right]\Big\}+\\
& (1-\pi)\left\{\rho\left[m(\overline{w}(R_f+(\overline{R}-R_f)\theta))+\left(m(\overline{w}(R_f+(\overline{R}-R_f)\theta))\right)-m(\underline{b})\right]+\right.\\
& (1-\rho)\left[m(\overline{w}(R_f-(R_f-\underline{R})\theta))+(m(\overline{w}(R_f-(R_f-\underline{R})\theta))-m(\underline{b}))\right]\Big\}
\end{aligned} \tag{3.11}$$

A necessary, and sufficient, condition for an interior optimum to exists is that $\frac{(\overline{R}-\underline{R})}{(R_f-\underline{R})} > \frac{(3+\lambda)-(\lambda-1)\rho}{4\rho}$ and is characterized by $\theta^{\overline{w},L_1}_{SPT} = Min\left\{\theta^*, \frac{R_f}{2(R_f-\underline{R})}\right\}$ where θ^* is characterized by

$$2\rho(\overline{R}-R_f)m'(\underline{w}(R_f+(\overline{R}-R_f)\theta^*)) - \frac{3+\lambda}{2}(1-\rho)(R_f-\underline{R})m'(\underline{w}(R_f-(R_f-\underline{R})\theta^*)) = 0 \tag{3.12}$$

then we can proof that

Proposition 3.4.7. *If $r'_R(x \mid m) \leq 0$, $m'(x) > 0$ and $m''(x) \leq 0$ then there is an $\overline{k} \in \mathbb{R}, \overline{k} < \infty$ such that $\theta^{\overline{w},L_1}_{SPT} = \theta^{\overline{w},L_0}_{SPT}$ if $\frac{(\overline{R}-\underline{R})}{(R_f-\underline{R})} \in [0, \frac{(3+\lambda)-(\lambda-1)\rho}{4\rho})$, $\theta^{\overline{w},L_1}_{SPT} > \theta^{\overline{w},L_0}_{SPT}$ for any $\frac{(\overline{R}-\underline{R})}{(R_f-\underline{R})} \in (\frac{(3+\lambda)-(\lambda-1)\rho}{4\rho}, \overline{k})$, and $\theta^{\overline{w},L_1}_{SPT} \leq \theta^{\overline{w},L_0}_{SPT}$ otherwise.*

However more interestingly is that we have that,

Proposition 3.4.8. *If $r'_R(x \mid m) \leq 0$ then $\theta^{\underline{w},L_1}_{SPT} > \theta^{\overline{w},L_1}_{SPT}$ for any $\frac{(\overline{R}-\underline{R})}{(R_f-\underline{R})} > \frac{2(1+\lambda)-(\lambda-1)\rho}{(\lambda+3)\rho}$ and $\theta^{\underline{w},L_1}_{SPT} = \theta^{\overline{w},L_1}_{SPT}$ otherwise.*

Which implies that among individuals facing the initial lottery for determining initial income, those who start with low income are more willing to undertake risk than those who got the high income. Intuitively, this is the case because a low income individual is in the loss domain, no matter what her invested amount is, with respect to the highest reference point. Therefore, such an individual will have incentives to increase her invested amount in order to reduce such loss feelings.

Summary SPT *(i) SPT predicts that those starting with low income, having lost the chance of getting the higher one, are more willing to invest in the risky alternative than those who "won" the chance to start with the high income level. (ii) In addition, SPT suggests that for those facing the lottery who got a high income, there are some values of $\overline{R}$ for which the average invested amount, over the support of ρ, is equal to the investment of high predetermined income individuals,*

that must hold for the arguments that maximize each functions, then we evaluate the value function in each region and the compare the level of both value functions to determine which one is greater.

3.4.4 Sampling–based reference points

Another possibility is that preferences are given by the average of $u\left(c|b_L^G\right)$ where b_L^G captures, in a single dimension, reference point beliefs an individual had when first focusing in the problem determined by $G_L(\cdot)$. Therefore, utility is given by

$$\theta_{SBPT}^{\omega,L} \in \arg\max_{\theta} \int u\left(y|b^{G_L}\right) dF\left(y \mid \theta\right). \tag{3.13}$$

Under this Sampling–based prospect theory (SBPT, Spiegler (2012)) framework we assume that $b^{G_L} = \int b \; dG_L(b)$ which captures somehow the recently held beliefs about reference points.

As in the previous SPT framework we know $\theta_{SBPT}^{\omega,L_0} = \theta_{PT}^{\omega,L_0}$ for any $\omega \in \{\underline{w}, \overline{w}\}$. However for those facing treatment $(\underline{w}, L_1)$ face the following problem

$$\begin{aligned} &\underset{0\leq\theta\leq1}{Max}\, \rho\left[m(\underline{w}(R_f+(\overline{R}-R_f)\theta)) - \lambda\left(m(b^G) - m(\underline{w}(R_f+(\overline{R}-R_f)\theta))\right)\right] + \\ &\qquad (1-\rho)\left[m(\underline{w}(R_f-(R_f-\underline{R})\theta)) - \lambda\left(m(b^G) - m(\underline{w}(R_f-(R_f-\underline{R})\theta))\right)\right] \end{aligned} \tag{3.14}$$

A necessary, and sufficient, condition for an interior optimum to exists is that $\frac{(\overline{R}-\underline{R})}{(R_f-\underline{R})} > \frac{1}{\rho}$ and characterized by

$$(1+\lambda)\rho(\overline{R}-R_f)m'(\underline{w}(R_f+(\overline{R}-R_f)\theta_{SBPT}^{\underline{w},L_1})) - (1+\lambda)(1-\rho)(R_f-\underline{R})m'(\underline{w}(R_f-(R_f-\underline{R})\theta_{SBPT}^{\underline{w},L_1})) = 0 \tag{3.15}$$

On the other hand those facing treatment $(\overline{w}, L_1)$ face the problem

$$\begin{aligned} &\underset{0\leq\theta\leq\frac{R_f}{4(R_f-\underline{R})}}{Max} \rho\left[m(\overline{w}(R_f+(\overline{R}-R_f)\theta)) + \left(m(\overline{w}(R_f+(\overline{R}-R_f)\theta)) - m(b^G)\right)\right] + \\ &\qquad (1-\rho)\left[m(\overline{w}(R_f-(R_f-\underline{R})\theta)) + \left(m(\overline{w}(R_f-(R_f-\underline{R})\theta)) + m(b^G)\right)\right] \end{aligned} \tag{3.16}$$

A necessary, and sufficient, condition for an interior optimum to exists is that $\frac{(\overline{R}-\underline{R})}{(R_f-\underline{R})} > \frac{1}{\rho}$ and is characterized by $\theta_{SBPT}^{\overline{w},L_1} = Min\left\{\theta^{**}, \frac{R_f}{4(R_f-\underline{R})}\right\}$ where θ^{**} is characterized by

$$2\rho(\overline{R}-R_f)m'(\underline{w}(R_f+(\overline{R}-R_f)\theta^{**})) - 2(1-\rho)(R_f-\underline{R})m'(\underline{w}(R_f-(R_f-\underline{R})\theta^{**})) = 0 \tag{3.17}$$

Henceforth we find that

Proposition 3.4.9. *If $r_R'(x \mid m) \leq 0$, $m'(x) > 0$ and $m''(x) \leq 0$ then there is an $\overline{k}' \in \mathbb{R}, \overline{k}' < \infty$ such that $\theta_{SBPT}^{\overline{w},L_1} = \theta_{SBPT}^{\overline{w},L_0}$ if $\frac{(\overline{R}-\underline{R})}{(R_f-\underline{R})} \in [0, \frac{1}{\rho})$, $\theta_{SBPT}^{\overline{w},L_1} > \theta_{SBPT}^{\overline{w},L_0}$ for any*

$\frac{(\overline{R}-\underline{R})}{(R_f-\underline{R})} \in (\frac{1}{\rho}, \overline{k})$, *and* $\theta^{\overline{w},L_1}_{SBPT} \leq \theta^{\overline{w},L_0}_{SBPT}$ *otherwise.*

the proof follows the same analysis of FOC as for proposition 3.4.7 where $\overline{k}'$ is such that $\theta^{\overline{w},L_0}_{SBPT}(\overline{k}') = 4(R_f - \underline{R})$. On the other hand, we find then that

Proposition 3.4.10. *If* $r'_R(x \mid m) \leq 0$, $m'(x) > 0$ *and* $m''(x) \leq 0$ *then there is an* $\underline{k} \in \mathbb{R}, \underline{k} < \infty$ *such that* $\theta^{\underline{w},L_1}_{SBPT} = \theta^{\overline{w},L_1}_{SPT}$ *for any* $\frac{(\overline{R}-\underline{R})}{(R_f-\underline{R})} \in [0, \underline{k})$, *and* $\theta^{\underline{w},L_1}_{SPT} > \theta^{\overline{w},L_1}_{SPT}$ *otherwise.*

The proof is similar proposition 3.7.1 where $\overline{k}''$ is such that $\theta^{\underline{w},L_1}_{SBPT}(\underline{k}) = 4(R_f - \underline{R})$

This happens mainly because they are not anymore subject to first–order risk aversion given that their reference point is sufficiently far, in absolute term, from their starting income and because we are assuming a piecewise linear gain–loss utility.

Summary SBPT *(i) SBPT does not predict any difference across individuals facing initial lottery. (ii) SBPT suggests that individuals facing a lottery for initial income who got high income invest strictly greater average amounts than predetermined high income individuals.*

Summary: Theoretical results

In Table 3.2 we summarize the predictions on the average invested amount on the risky alternative with a specific CRRA function, $m(c) = \frac{c^{1-\gamma}}{1-\gamma}$, assuming that $\overline{R}$ is a continuous random variable distributed uniformly between 1.05 and 1.3 and ρ also uniformly distributed between 0.1 and 0.9. Results hold for values of $\lambda > 1$ and $\gamma \in (0, 1)$.

Table 3.2: Theoretical Results‡ on average invested levels $\overline{\theta}^{\omega,L}_M$.

	EUT*	PT**	SPT†	SBPT††
RV effect				
$\overline{\theta}^{\underline{w},L_1} - \overline{\theta}^{\underline{w},L_0}$	$= 0$	$= 0$	> 0	> 0
$\overline{\theta}^{\overline{w},L_1} - \overline{\theta}^{\overline{w},L_0}$	$= 0$	$= 0$	$\lesseqgtr 0$	> 0
IV effect				
$\overline{\theta}^{\overline{w},L_0} - \overline{\theta}^{\underline{w},L_0}$	$= 0$	$= 0$	$= 0$	$= 0$
$\overline{\theta}^{\overline{w},L_1} - \overline{\theta}^{\underline{w},L_1}$	$= 0$	$= 0$	≤ 0	$= 0$

‡ Common assumptions: $m' > 0, m'' \leq 0$ and $r'_R(\cdot) = 0$.

* Propositions 3.4.1-3.4.3, ** props. 3.4.4-3.4.5,

† props. 3.4.6-3.7.1, †† props. 3.4.9-3.4.10

Hence, if, across individuals with same income level ω, variation L_1 has a positive effect on the invested amount with respect to variation L_0, we can favor the set of theories where recently held expectations determine a reference point (i.e. SPT and SBPT) as opposed to those where past expectations do not affect them (i.e. EUT and PT). If that is indeed the case, we further can distinguish between SPT and SBPT: SPT is the only theory, among those considered, predicting that under reference point variation L_1, those who start with low income ($\omega = \underline{w}$) are more willing to undertake risk, investing larger amounts, than those who got a high income ($\omega = \overline{w}$). SPT also suggests that for those with high income, the difference on the invested amount between those who face the initial lottery and those who did not is somewhat small, rather than a large difference as suggested by SBPT.

In general, all theories suggest that the higher the success probability is, the higher the invested amount in the risky alternative should be. They also predict that larger average rate of return make individuals more likely to invest strictly positive amounts of their income. SPT and SBPT predict as well that individuals facing initial lottery ($L = L_1$) are more willing to invest than those starting with a predetermined income ($L = L_0$) as opposed to EUT and PT, which predict no difference. However, only SPT suggests that those individuals with low income and initial lottery ($\underline{w}, L_1$) are more willing to invest than those with high income and initial lottery ($\overline{w}, L_1$).

Only the last two frameworks are able to produce differences across reference dependent variations. Finally, notice that any forward looking type of expectation–based reference points are held constant across treatments and therefore can't affect the observed decisions. In that sense an EBPT à la Köszegi & Rabin (2006) cannot have any observational implication in differentiating among treatments.

In the following section we include some descriptive statistics of the sample.

3.5 Experimental Findings

Table 3.3 depicts how successful the randomization was across treatments. We see that the randomization of the IV was balanced for female ratio, age, economic strata, percentage of people with higher education, last month income (in 2008 USD dollars) and household income (in number of minimum wages). The RV was also successful in every category apart from the percentage of people with higher education degrees, there is a significant difference but only at a 10% level. Given this, we control for individual characteristics in our econometric specification.

Table 3.3: Balanced Sample across treatments

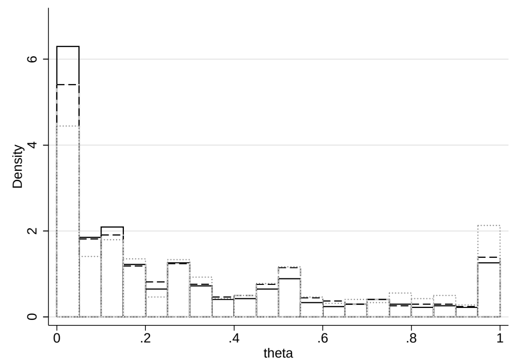
	Control			Treatment			$H_0 : 1-2=0$	
	N	Mean (1)	SE	N	Mean (2)	SE	t-stat	pval
Income Variation, IV								
Female	108	0.64	0.05	108	0.66	0.05	0.28	0.78
Age	107	43.48	1.25	107	41.79	1.25	-0.96	0.34
Econ. strata†	107	2.64	0.07	108	2.56	0.07	-0.83	0.41
Higher educ	108	0.40	0.05	108	0.48	0.05	1.23	0.22
Last month income	105	362.24	27.20	107	300.91	26.94	-1.60	0.11
HH income	108	5.68	0.28	107	5.08	0.28	-1.47	0.14
Reference Variation, RV								
Female	150	0.61	0.04	66	0.73	0.06	1.62	0.11
Age	149	42.78	1.06	65	42.29	1.60	-0.25	0.80
Econ. strata†	149	2.62	0.06	66	2.53	0.09	-0.90	0.37
Higher educ	150	0.48	0.04	66	0.35	0.06	-1.80	0.07*
Last month income	146	322.27	23.18	66	351.23	34.48	0.70	0.49
HH income	149	5.31	0.24	66	5.55	0.36	0.54	0.59

† Official Colombian system of classification by economic status. It takes values between 1 and 6. Lower values translate into more deprived physical households. * $p < 0.1$, ** $p < 0.5$, *** $p < 0.01$. notes: Differences in total observations are due to information not properly reported

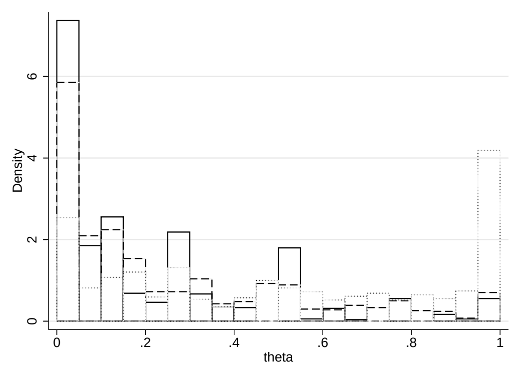
In the next section we provide some descriptive results from individual decisions.

3.5.1 Descriptive statistics

First, consider the distribution of percentage of invested income found in the data (i.e. $\theta_{i,t,s}$). Figure 3.2 shows on Panel (a) the density of this variable distinguishing across different rate of returns if successful (i.e, $\overline{r}_s$). One can easily notice that as the probability of success increases the density becomes more negatively skewed implying higher invested amounts as the likelihood of a successful outcome becomes larger. On the other hand Panel (b) depicts the density of the same variable but varying the success probabilities (i.e. ρ^t), again we notice that the bulk of the density moves right as the rate of return increases.[7]



(a) By rate or return, $\overline{r}_1 = .05$ (solid), $\overline{r}_3 = .15$ (dashed), $\overline{r}_3 = .3$ (dotted)



(b) By success probability. $\rho^1 = .1$ (solid), $\rho^3 = .5$ (dashed), $\rho^5 = .9$ (dotted)

Figure 3.2: Density of invested Income

It is worth noting that in both panels there is a considerable concentration of values at $\theta_{t,s} = 0$ and a bounded outcome between $[0, 1]$ which contravenes the standard

[7]In the appendix 3.7.2 Figures 3.6 and 3.5 provide histograms varying both success probability and rate of return.

statistical assumptions of the linear model, suggesting the use of a mixed discrete-continuous distribution to analyse the data. We will come back to this issue below in the econometric specification.

A preliminary analysis of the data is depicted in Figure 3.3 which shows the fitted mean from a fractional polynomial regression which takes the following variables as dependent ones. In Panel (a) it takes the average percentage of income invested across different investment alternatives at a given round, that is $\overline{\theta}_{i,t} = (1/5)\sum_{s=1}^{5}\theta_{t,i,s}$, which we can associate to changes in ρ^t across $t \in \{1, \cdots, 5\}$; while in Panel (b) we take the average percentage of invested income in an alternative across rounds, that is $\overline{\theta}_{i,s} = (1/5)\sum_{t=1}^{5}\theta_{i,t,s}$ which can be associated with changes to $\overline{r}_s$ for any investment alternative $s \in \{1, \cdots, 5\}$.[8]

We show solid line for individuals under treatments $(\overline{w}, L_1)$, dotted line represents individuals starting with $(\underline{w}, L_1)$ treatment, dashed lines represent high starting income individuals but which did not face initial lottery $(\overline{w}, L_0)$ and finally the long dashed–dotted line represents low income and no initial lottery $(\underline{w}, L_0)$.[9]

From both graphs in Figure 3.3 it is ready noticeable, though statistically significance is postponed until next session, that treatments (w, L) have an effect on individuals decisions. In particular, considering fixed reference point treatment, we notice that those with higher starting income are investing slightly larger amounts than those starting with lower income. This would be small evidence in favour of decreasing relative risk aversion if it happens to be statistically significant, or in favour of constant relative risk aversion if it is not.

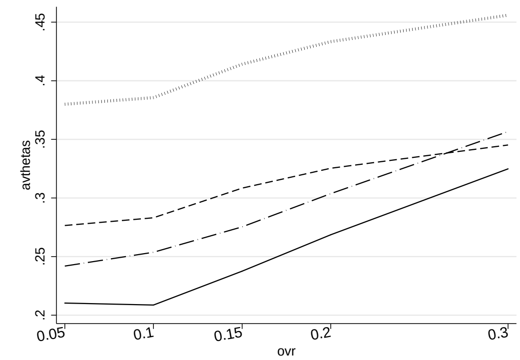
Importantly, notice that those who face the initial lottery and got a high income, are generally less willing to invest larger amounts, at any success probability or rate of return, than those who started with a fixed high income. This result, if statistically significant, contradicts our theoretical approach for SPT.

On the other hand, individuals that got a low income after the initial lottery are eager to invest larger amounts, than those individuals starting with a low income without facing the initial lottery. A result clearly in line with the behavioural implications from our SPT model.

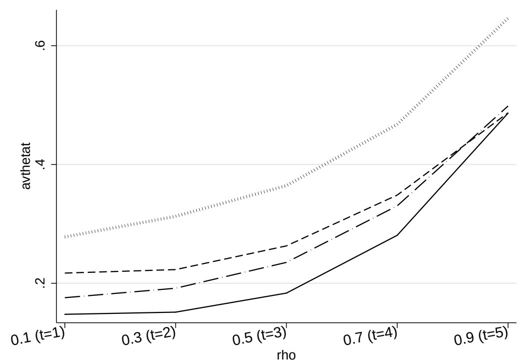
Neither EUT nor PT predict these additional differences across treatments (ω, L). On the other hand, SPT and $SBPT$ do predict differences across treatments, as expressed in Table 3.2. In the following section we present the econometric analysis which

[8] See in appendix 3.7.2, Figures 3.7 and 3.8 which distinguishes $\theta_{t,s}$ among all success probabilities and rate of returns without taking averages. The patterns are nonetheless equal to the figure presented above.

[9] See figure 3.4 in the appendix for the average of those who chose to change investment with the fair insurance (i.e. $D_{i,t,s} = 1$). As we notice it is generally less than 30% of the population who follow this decision. The study of such behaviour is left for future work.



(a) By rate or return



(b) By success probability

Figure 3.3: Average percentage of invested income

Source: solid $(\overline{w}, L_1)$, dotted $(\underline{w}, L_1)$, dashed $(\overline{w}, L_0)$,dashed–dotted $(\underline{w}, L_0)$

elucidate which of the considered theories on formation of reference points is more likely to explain the qualitative results.

3.5.2 Econometric Specification

Our specification takes as dependent variable the percentage invested by individual i in state s given success probability in round t (i.e. $\theta_{i,t,s}$). Our main econometric approximation controls for experimental treatments $\mathbf{1}_i[\underline{w}, L_0]$,$\mathbf{1}_i[\overline{w}, L_0]$, $\mathbf{1}_i[\underline{w}, L_1]$, $\mathbf{1}_i[\overline{w}, L_1]$. We include fixed effects for investment alternative s (d_s) and by success probability or round t (d_t).

In our analysis we first provide OLS and Tobit estimates as benchmarks. However, Cook, Kieschnick & McCullough (2008) identify three specification errors when following the linear approach whenever dependent variable lies within the $[0, 1]$ interval: First, the conditional expectation of a continuous percentage is only defined on the interval $[0, 1]$. Therefore, it must be a non–linear function of the explanatory variables. Second, the conditional variance and the conditional mean must be collinear, because as the later approaches either boundary the conditional variance must change accordingly. And third, different decision processes determine whether and individual is willing to invest or not (i.e. $\mathbb{1}[\theta_{i,t,s} = 0]$) and how much she is willing to invest (i.e. optimal level $\theta_{i,t,s}$).

On the same token defining a Tobit on a percentage, to deal with the first two problems, may not be an appropriate strategy if the observed data is not censored by nature but, rather, values outside the $[0, 1]$ interval are not feasible for such type of data.[10] If that is the case, a mixed discrete-continuous distribution, given there are mass points at zero and one (see Figure 3.2) may be a better statistical approximation.

We therefore deal with the statistical problems posed above, estimating a Generalised Linear Model equivalent to the quasi–likelihood model developed by Papke & Wooldridge (1996) (GLM) and a Zero–and–One Inflated Beta model (ZOIB) following Ospina & Ferrari (2010, 2011). This later econometric approach has two advantages over GLM: firstly, we do not need to use asymptotic properties raised by quasi–likelihood approach to validate its results and secondly, it allows for decision $\mathbb{1}[\theta_{i,t,s} = 0]$ to be generated by a different process than percentage $\theta_{i,t,s} > 0$.[11]

Consider equation 3.18 providing the linear approximation of the estimated model.

$$\theta_{i,t,s} = a_1 \mathbf{1}_i[\overline{w}, L_0] + a_2 \mathbf{1}_i[\underline{w}, L_1] + a_3 \mathbf{1}_i[\overline{w}, L_1] + \eta_{i,t,s}, \tag{3.18}$$

where $\eta_{i,t,s} = a_0 + \sum_{t=1}^{5} \gamma_t d_t + \sum_{s=1}^{5} \delta_s d_s + z_i \delta + \epsilon_{i,t,s}$. Being $\epsilon_{i,t,s}$ orthogonal to the covariates z_i, given the randomization of the treatments, and assumed to be correlated at the individual level. Replacing back $\eta_{i,t,s}$ in equation 3.18, let us define this data generating model as

$$\theta_{i,t,s} = \mathbf{a}' X_{i,t,s} + \epsilon_{i,t,s},$$

where $X_{i,t,s}$ is a column vector containing the regressors with a constant term as its first argument.

Notice first that OLS specification 3.18 has a close relation with the average of the

[10] However, in our present set–up, θ are actually censored given that individuals in the experiment do not have access to outside credit markets for an experimental session.

[11] Analytically this difference has a mathematical link with decision problems in section 3.4. The equation behind $\mathbb{1}[\theta_{i,t,s} = 0]$ is related to necessary conditions for a corner solution, usually the sign of the FOC and SOC. While the expression for percentage $\theta_{i,t,s}$ has to do with the FOC holding with equality.

invested percentages, $\overline{\theta}^{\omega,L}$ (associated with predictions from section 3.4) in particular,[12]

$$
\begin{aligned}
\hat{\mathbb{E}}[\theta_{i,t,s} \mid \mathbf{1}_i[\underline{\omega}, L_0] = 1, \cdot] &= \hat{a}_0 & \xrightarrow{p} \overline{\theta}^{\underline{\omega},L_0}, \\
\hat{\mathbb{E}}[\theta_{i,t,s} \mid \mathbf{1}_i[\underline{\omega}, L_1] = 1, \cdot] &= \hat{a}_0 + \hat{a}_2 & \xrightarrow{p} \overline{\theta}^{\underline{\omega},L_1}, \\
\hat{\mathbb{E}}[\theta_{i,t,s} \mid \mathbf{1}_i[\overline{\omega}, L_0] = 1, \cdot] &= \hat{a}_0 + \hat{a}_1, & \xrightarrow{p} \overline{\theta}^{\overline{\omega},L_0}, \\
\hat{\mathbb{E}}[\theta_{i,t,s} \mid \mathbf{1}_i[\overline{\omega}, L_1] = 1, \cdot] &= \hat{a}_0 + \hat{a}_1 + \hat{a}_2 + \hat{a}_3 & \xrightarrow{p} \overline{\theta}^{\overline{\omega},L_1}.
\end{aligned}
$$

For the OLS specification $\epsilon_{i,t,s}$ is assumed to be normally distributed while for the Tobit case is assumed to present a truncated distribution at 0 and 1. On the other hand, for the ZOIB specification, we follow Ospina & Ferrari (2010), and define that $\epsilon_{i,t,s}$ is distributed with a mixed discrete–continuous density function $h(\cdot)$ such that,

$$
h(\theta_{i,t,s}; \pi_0, \pi_1, \mu, \phi) = \begin{cases} \pi_0 & \text{if} \quad \theta_{i,t,s} = 0, \\ (1-\pi_0)(1-\pi_1)\mathfrak{b}(\theta_{i,t,s}; \mu, \phi) & \text{if} \quad 0 < \theta_{i,t,s} < 1, \\ \pi_1 & \text{if} \quad \theta_{i,t,s} = 1. \end{cases} \tag{3.19}
$$

Where $\pi_0 = \frac{\exp(\mathbf{a}_0' X_{0i,t,s})}{1+\exp(\mathbf{a}_0' X_{0i,t,s})}$, $\pi_1 = \frac{\exp(\mathbf{a}_1' X_{1i,t,s})}{1+\exp(\mathbf{a}_1' X_{1i,t,s})}$. And $\mathfrak{b}(\cdot)$ is the density function of a beta distribution with mean equation $0 < \mu = \frac{\exp(\mathbf{a}' X_{i,t,s})}{1+\exp(\mathbf{a}' X_{i,t,s})} < 1$ and precision $\phi \equiv \exp(\beta)$, denoted by $\mathfrak{B}(\mu, \phi)$, such that

$$
\mathfrak{b}(\theta_{i,t,s}; \mu, \phi) = \frac{\Gamma(\phi)}{\Gamma(\mu\phi)\Gamma(\phi(1-\mu))} \theta_{i,t,s}^{\mu\phi-1} \theta_{i,t,s}^{\phi(1-\mu)-1}, \tag{3.20}
$$

where $\Gamma(\cdot)$ is the gamma function. The beta distribution has been shown suitable for modelling dependent variables that belong to the $(0,1)$ interval. As we study different processes governing the decisions to invest and how much to invest, we allow estimates for $\mathbf{a}_0$ to differ from $\mathbf{a}_1$ and $\mathbf{a}$. However, we specify covariates for the zero cases to be the same as those for strictly positive values (i.e. $X_{0i,t,s} = X_{i,t,s}$) and for the one cases we include an additional constant $X_{1i,t,s} = (1, 0, \cdots, 0)$.

Therefore, the log–likelihood function implied by the ZOIB model is given by

$$
L(\kappa) = \frac{1}{N} \sum_{i=1}^{n} \sum_{t=1}^{5} \sum_{s=1}^{5} \ln \left[\pi_0^{\mathbb{1}[\theta_{i,t,s}=0]} \pi_1^{\mathbb{1}[\theta_{i,t,s}=1]} \left((1-\pi_0)(1-\pi_1)\mathfrak{b}(\theta_{i,t,s}) \right)^{\mathbb{1}[0<\theta_{i,t,s}<1]} \right], \tag{3.21}
$$

where $\kappa = (\mathbf{a}, \mathbf{a}_0, a_1, \beta)$ is the parameter vector to be estimated.

[12] Where $\hat{\mathbb{E}}$ is the empirical expectation once we discount the investment alternative and round fixed effects.

Even though the nature of sequential decisions by subjects suggests a dynamic panel data model, where we explicitly need to account for the time dependence of the error term, the following understanding allows us to consider a model in which we only need to include rounds dummy and still get consistent estimators. The problem of initial conditions arises in dynamic panel data models with unobserved heterogeneity or serially correlated shocks, therefore we must account for the initial conditions whenever they are endogenous. However, the random assignment of treatments imply that initial conditions are indeed exogenous, meaning that initial income is fixed and it was not generated by the process generating the sample of subsequent income levels. We do however allow for robust clustered variance at the individual level.

On the same token, we consider that the experimental setup of not presenting for how many rounds individuals were going to take decisions for, allows us to follow this approach instead of following dynamic setting where future value functions are to be solved first by backward induction.

3.5.3 Results

The estimated coefficients using OLS, TOBIT, GLM and ZOIB specifications are given in Table 3.4. The last two columns include the prospect theory parameters, recovered in the first stage, into the z_i variables.

Results across specifications suggest that individuals with high income levels and fixed reference points invest larger amounts than those with low incomes and fixed reference points. However, this coefficient is not statistically significant suggesting a constant relative risk aversion in income.

Nevertheless, coefficients from every specification, suggest that those with larger initial income, and predetermined reference point, are more willing to invest a strictly positive amount than those with lower income.

The way in which the initial income is allocated seems to have a strong and persistent effect on individual decisions. In particular, low income individuals who faced the coin flip are more willing to invest (they are less likely to invest zero in the risky alternative, see zero inflate coefficients). They are also investing larger amounts, than those who did not face the initial lottery (main coefficients). Finally, high income individuals facing initial lottery tend to shy away from risky investments more often than low income individuals facing the same initial lottery.

On the other hand, coefficients recovered from the first stage have consistent signs: the more risk loving an individual is (i.e. $1-\gamma$ larger) the larger the invested amount would be while more loss averse individuals (i.e. larger λ) invest significantly lower amounts.

In the last column we also include an indicator variable that captures whenever an individual has experienced a loss in the previous round. What we observed is that such experience make individuals more cautious and less willing to invest. However, the effect of our treatments remain. This is reassuring of our intuition that the treatment effect is mediated through individual preferences, by affecting the reference point, and not through beliefs in the shape of winning–losing streaks. The sign goes in the opposite direction than the one predicted by Köszegi & Rabin (2007) which claim that risk–loving behaviour is expected after an agent have, unexpectedly, lost money. This suggest then that subjects in our experiment are actually aware of the chances of losing money.

Table 3.4: Estimated Coefficients

main: a	OLS	TOBIT	GLM	ZOIB	ZOIB+S1	ZOIB+S1+L
	(1)	(2)	(3)	(4)	(5)	(6)
$\mathbf{1}[\overline{w}, L_0]$	0.021	0.058	0.111	-0.220	-0.190	-0.212
	(0.040)	(0.050)	(0.207)	(0.168)	(0.176)	(0.179)
$\mathbf{1}[\underline{w}, L_1]$	0.128**	0.163**	0.610**	0.417**	0.431**	0.424*
	(0.052)	(0.065)	(0.245)	(0.245)	(0.210)	(0.229)
$\mathbf{1}[\overline{w}, L_1]$	-0.185***	-0.258***	-0.919***	-0.262	-0.253	-0.202
	(0.069)	(0.086)	(0.345)	(0.263)	(0.278)	(0.292)
Constant	0.141***	0.048	-1.708***	-1.054***	-0.778***	-0.568*
	(0.026)	(0.036)	(0.154)	(0.116)	(0.338)	(0.343)
$1-\gamma$					0.076	0.095
					(0.252)	(0.263)
λ					-0.040*	-0.042**
					(0.020)	(0.021)
α					-0.246	-0.285
					(0.454)	(0.455)
$Loss_{t-1}$						-0.194**
one inflate: a_1						
Constant				-2.471***	-2.504***	-2.388***
zero inflate: $\mathbf{a}_0$				(0.190)	(0.195)	(0.189)
$\mathbf{1}[\overline{w}, L_0]$				-1.077***	-0.992***	-1.111***
				(0.352)	(0.359)	(0.381)
$\mathbf{1}[\underline{w}, L_1]$				-0.704*	-0.723	-0.911*
				(0.424)	(0.457)	(0.502)
$\mathbf{1}[\overline{w}, L_1]$				1.635***	1.578***	1.952***
				(0.595)	(0.629)	(0.685)
Constant				-0.226	-0.358	-2.779***
				(0.229)	(0.748)	(0.900)
(γ, λ, π)					yes	yes
$Loss_{t-1}$						2.230***
$ln(\phi) = \beta$				0.843***	0.879***	0.831***

Obs: 5400 (before–last 4950, last 3960), Clustered s.e. round and investment fe. γ risk aversion, λ loss aversion, α probability weight * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Even though coefficients are suggestive of the importance of the reference point determination, in Table 3.5 we provide average marginal effects of the econometric approximations of columns (2)-(4). We can compare those marginal effect with the theoretical predictions in table 3.2 from the previous section. We stress that EUT and PT do not predict any difference across RV.

Firstly, marginal effects also suggest a constant relative risk aversion: among individuals with predetermined initial income treatment, those with larger initial income invest statistically the same percentage as their lower income counterparts. We cannot reject the null hypothesis $H_0 : \theta^{\overline{w},L_0} - \theta^{\underline{w},L_0} = 0$. This renders validity to our theoretical approach of CRRA preferences.

More importantly, focusing in the ZIB model, we find strong evidence of the relevance of reference–points formation. First, individuals with low initial income facing the stochastic determinacy of their reference point (in other words that had the chance to start with a high income but lost it) are more prone to invest, as much as 10.3% more, than individuals with a predetermined low income ($H_0 : \theta^{\underline{w},L_1} - \theta^{\underline{w},L_0} = 0$ is rejected). This is consistent with behavioural prediction from SPT and SBPT and inconsistent with EUT and PT.

Table 3.5: Estimated Average Marginal Effects

	TOBIT	GLM	ZIB
RV effect			
(1) $H_0 : \overline{\theta}^{\underline{w},L_1} - \overline{\theta}^{\underline{w},L_0} = 0$	0.163**	0.128**	0.103**
	(0.065)	(0.052)	(0.044)
(2) $H_0 : \overline{\theta}^{\overline{w},L_1} - \overline{\theta}^{\overline{w},L_0} = 0$	-0.095*	-0.058	-0.002
	(0.055)	(0.045)	(0.030)
IV effect			
(3) $H_0 : \overline{\theta}^{\overline{w},L_0} - \overline{\theta}^{\underline{w},L_0} = 0$	0.058	0.021	-0.002
	(0.050)	(0.040)	(0.029)
(4) $H_0 : \overline{\theta}^{\overline{w},L_1} - \overline{\theta}^{\underline{w},L_1} = 0$	-0.200***	-0.164***	-0.106**
	(0.070)	(0.056)	(0.044)
Observations	5400	5400	5400

Clustered standard errors in parentheses. Includes round and investment alternative dummy * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

On the other hand, the negative sign in the second row (corresponding to $H_0 : \theta^{\overline{w},L_1} - \theta^{\overline{w},L_0}$) suggests that initial lottery on reference points has an asymmetric effect between those who started winning and those losing in a inconsistent manner with stochastic reference points. This would suggest a sort of break–even and satisfaction effects: those who start losing feel that they have lost already and they are willing to undertake more risk in order to reach the level they could have attained; on the other hand, those that won the high initial income are already satisfied for not having lost and thus, are less willing to undertake unnecessary risk. However, this difference is not statistically significant different from zero.

The difference in signs of the RV effect on both income levels (rows (1) vs (2)) suggests that the statistically significant difference on the proportion of individuals with a higher degree between samples L_0 and L_1, is not the driving force behind the RV effect. If that would be the case, we do not see why the effect of RV over $\omega = \underline{w}$ and $\omega = \overline{w}$ treatments should have different signs.

Finally, we find that willingness to invest positive amounts from low income individuals who faced the initial lottery is statistically larger than high income also facing the initial lottery. We reject $H_0 : \theta^{\overline{w},L_1} - \theta^{\underline{w},L_1} = 0$ even though individuals seem to present constant relative risk aversion. Among the considered theories, SPT is the only one which predicts such effect.

It is also noteworthy that Tobit specification gives upward biased estimates of the real effect. Once we correct the specification problems explained above, we notice the estimated average marginal effect with ZOIB are lower than previous Tobit estimates. In particular, the effect of the initial lottery on high income individuals is not statistically significant.

Our results differ from the empirical ones of Gomes (2005). He finds that individuals facing good news will demand more from risky assets than those who faced bad news. If past expectations play a role in determining the reference point our results suggest the effect should reversed. Our experimental results are also somehow different from Post et al. (2008) because they find that a similar disposition effect, more willingness to undertake risk, between individuals who experienced bad news and those who experienced good news early in the game. The later difference could be explained by the nature of their forward–looking definition of reference points as opposed to the backward–looking one suggested here.

A final piece of evidence on multiple reference points is provided in figure 3.9. The kernel bounded densities of invested percentage across rounds suggest that individuals under $L_1, \underline{w}$ are bunching together close to one additional reference point compared to those under treatment $L_0, \underline{w}$.[13]

[13] Where we use the Reflection estimator at the boundaries 0 and 1, such that $\hat{f}^r(\theta; h, 0, 1) = \frac{1}{\sum_i^n w_i} \sum_i^n \frac{w_i}{h} K^r(\theta; \theta_i, h, 0, 1)$ where $\theta \in [0, 1]$ and $K^r(\theta; \theta_i, h, 0, 1) = K(\frac{\theta - \theta_i}{h}) + K(\frac{\theta + \theta_i}{h}) + K(\frac{\theta + \theta_i - 2}{h})$.

3.6 Conclusion

We provide artefactual field experiment evidence on reference–dependent preferences and investment decisions on a sample of vulnerable small entrepreneurs. In particular we find that having a high initial income increases individual's likelihood to invest a positive amount if and only if they did not face the lottery to determine their reference points. Contrary to this, agents who started with a low income were more prone to invest (even at lower rates of return) if they started with uncertainty about their reference points. In this sense, having unfulfilled recent beliefs about a reference point triggers risk–loving behaviour only on the low income individuals while promoting risk–averse behaviours in high–income individuals, only the former effect is captured by the simple theoretical model provided.

Several authors have provided arguments on how expectations on future events can determine reference points. Our experimental evidence suggests that past expectations could affect also reference points determination in close relatedness to multiple reference points suggested by Köszegi & Rabin (2007). We provided a model of stochastic reference points that replicates qualitatively most of the experimental results.

3.7 Appendix

3.7.1 Proofs

Proof of prop 3.4.2. If there exists an interior solution, the solution is characterized by equation 3.4 with equality and negative sign on the SOC condition. By the Implicit Function Theorem and replacing $\overline{R} - R_f = \overline{r}$ and $R_f - \underline{R} = \underline{r}$, we get

$$\frac{\partial \theta^{\omega,L}_{EUT}}{\partial \omega} = \frac{(1-\rho)\underline{r}m''(\omega(1-\theta\underline{r}))\omega(1-\theta\underline{r}) - \rho\overline{r}m''(\omega(1+\theta\overline{r}))\omega(1+\theta\overline{r})}{\rho\overline{r}^2 m''(\omega(1+\theta\overline{r})) + (1-\rho)\underline{r}^2 m''(\omega(1-\theta\underline{r}))} \tag{3.22}$$

Knowing that the coefficient of relative risk aversion is given by $r_R(x \mid m) = -\frac{m''(x)}{m'(x)}x$ then we can show that if it is non increasing $r'_R(x \mid m) \leq 0$ then both numerator and denominator of equation 3.22 are negative. The equality is given whenever there is a corner solution at zero. □

Proof of 3.4.6. We already know that for any $\frac{(\overline{R}-\underline{R})}{(R_f-\underline{R})} < \frac{(1+\lambda)-(\lambda-1)\rho}{2\rho}$ we have $\theta^{\underline{w},L_0}_{SPT} = 0$. It is easy to show that $\frac{(1+\lambda)-(\lambda-1)\rho}{2\rho} > \frac{2(1+\lambda)-(\lambda-1)\rho}{(3+\lambda)\rho}$ whenever $\lambda > 1$. Therefore, if $\frac{(\overline{R}-\underline{R})}{(R_f-\underline{R})} \leq \frac{2(1+\lambda)-(\lambda-1)\rho}{(3+\lambda)\rho}$ then $\theta^{\underline{w},L_1}_{SPT} = \theta^{\underline{w},L_0}_{SPT} = 0$ but when $\frac{(\overline{R}-\underline{R})}{(R_f-\underline{R})} > \frac{2(1+\lambda)-(\lambda-1)\rho}{(3+\lambda)\rho}$

we have that $\theta^{\underline{w},L_1}_{SPT}$ is defined implicitly by equation 3.10 which evaluated in LHS of expression 3.6 gives us a negative value, implying then that $\theta^{\underline{w},L_1}_{SPT} > \theta^{\underline{w},L_0}_{SPT}$. □

Proof of prop 3.4.7. Let us denote Denote $\overline{k}$ the value of $\frac{\overline{R}-\underline{R}}{R_f-\underline{R}}$ such that $\theta^{\overline{w},L_0}_{SPT}(\overline{k}) = \frac{R_f}{2(R_f-\underline{R})}$. The first equality follows because there is a corner solution equal to zero for $\theta^{\overline{w},L_1}_{SPT}$ whenever $\frac{(\overline{R}-\underline{R})}{(R_f-\underline{R})} \leq \frac{(3+\lambda)-(\lambda-1)\rho}{4\rho}$ and for $\theta^{\overline{w},L_0}_{SPT}$ whenever $\frac{(\overline{R}-\underline{R})}{(R_f-\underline{R})} \leq \frac{(1+\lambda)-(\lambda-1)\rho}{2\rho}$, which is characterized by equation 3.7. Given $\lambda > 1$ we know that $\frac{(3+\lambda)-(\lambda-1)\rho}{4\rho} < \frac{(1+\lambda)-(\lambda-1)\rho}{2\rho}$. The second inequality follows because by replacing θ^* implied by equation 3.15 into equation 3.6 gives us a negative sign implying a lower optimal $\theta^{\overline{w},L_o}_{SPT}$. Finally, the last inequality is due to $\theta^{\overline{w},L_1}_{SPT}$ being bounded above by $\frac{R_f}{2(R_f-\underline{R})}$ given the incentive to avoid losses, while $\theta^{\overline{w},L_0}_{SPT}$ is bounded above by 1. □

Proof of prop . Combine proposition 3.4.6 and 3.4.7 to get the result. □

3.7.2 Further descriptives

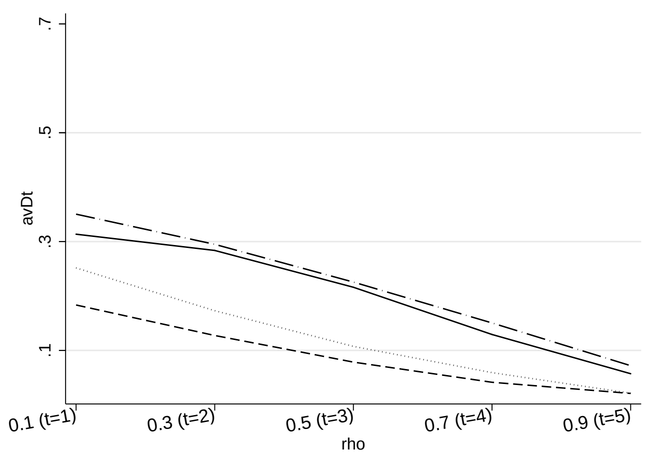
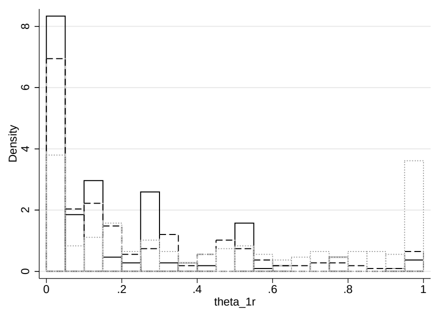
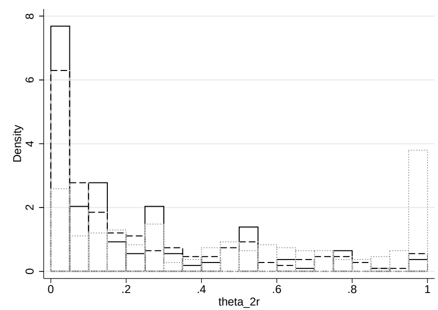


Figure 3.4: Percentage of Insurance by Rounds

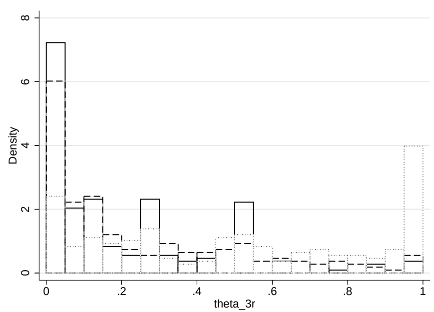
Source: solid $(\overline{w}, L_1)$, dotted $(\underline{w}, L_1)$, dashed $(\overline{w}, L_0)$,dashed–dotted $(\underline{w}, L_0)$



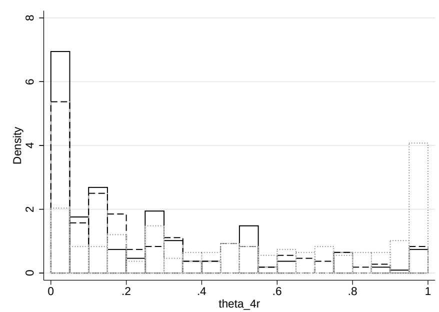
(a) $\overline{r}_1 = 0.05$



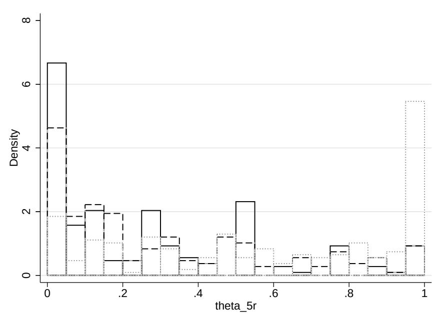
(b) $\overline{r}_2 = 0.1$



(c) $\overline{r}_3 = 0.15$



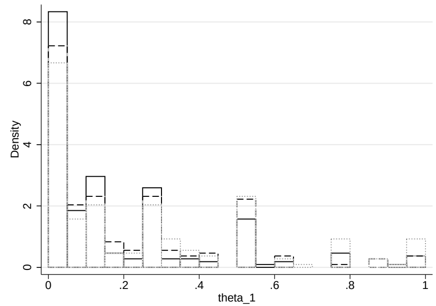
(d) $\overline{r}_4 = 0.2$



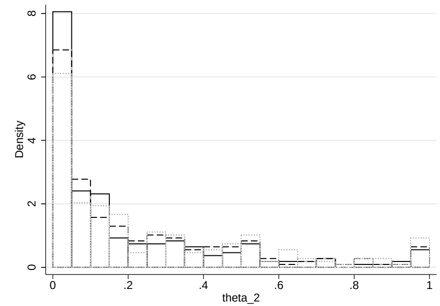
(e) $\overline{r}_5 = 0.3$

Figure 3.5: Density of Income Invested by option

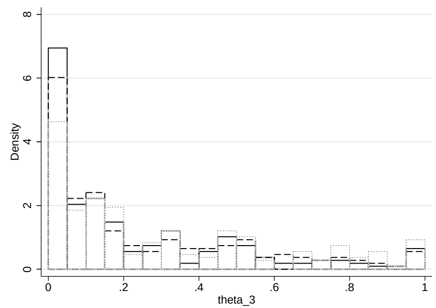
Source: $\rho^1 = 0.1$ (solid), $\rho^3 = 0.5$ (dashed), $\rho^5 = 0.9$ (dotted)



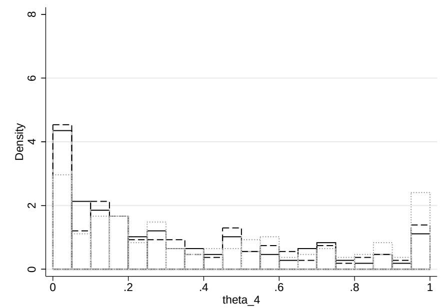
(a) $\rho^1 = 0.1$



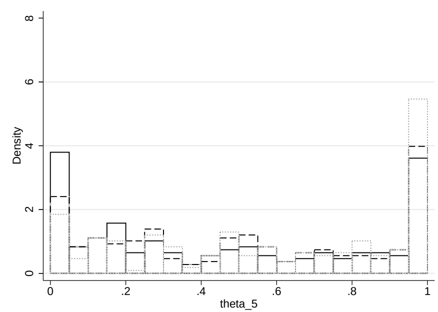
(b) $\rho^2 = 0.3$



(c) $\rho^3 = 0.5$



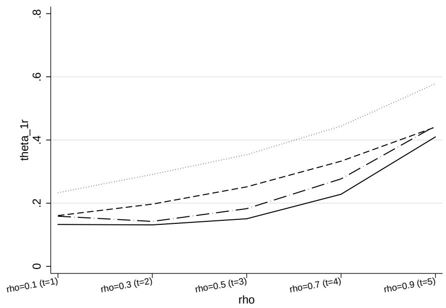
(d) $\rho^4 = 0.7$



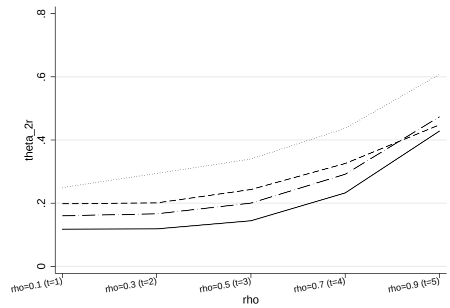
(e) $\rho^5 = 0.9$

Figure 3.6: Density of Income Invested by rounds

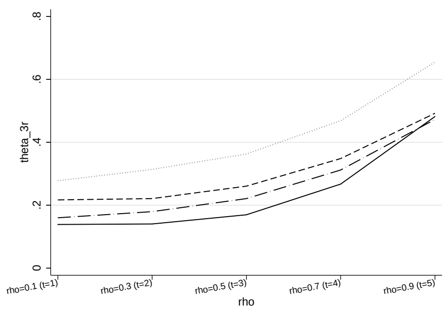
Source: $\overline{r}_1 = 0.05$ (solid), $\overline{r}_3 = 0.15$ (dashed), $\overline{r}_5 = 0.3$ (dotted)



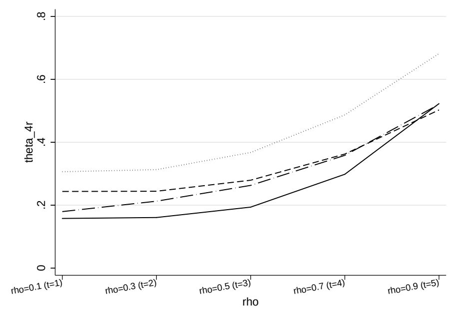
(a) $\overline{r}_1 = 0.05$



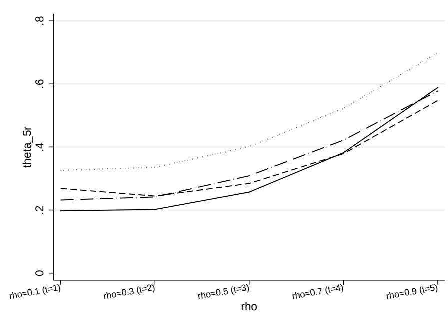
(b) $\overline{r}_2 = 0.10$



(c) $\overline{r}_3 = 0.15$



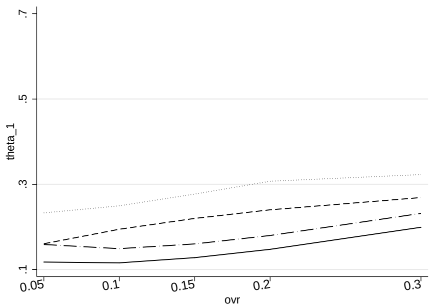
(d) $\overline{r}_4 = 0.20$



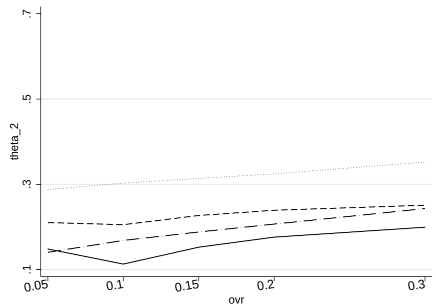
(e) $\overline{r}_5 = 0.30$

Figure 3.7: Average Percentage of Income Invested by Rounds

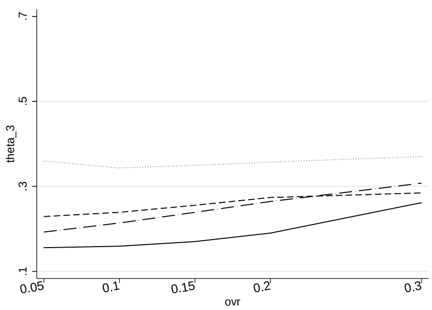
Source: solid $(\overline{w}, L_1)$, dotted $(\underline{w}, L_1)$, dashed $(\overline{w}, L_0)$,dashed–dotted $(\underline{w}, L_0)$



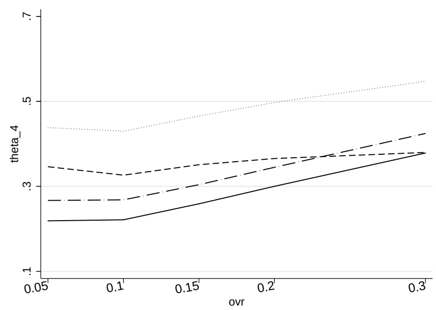
(a) $\rho^1 = 0.1$



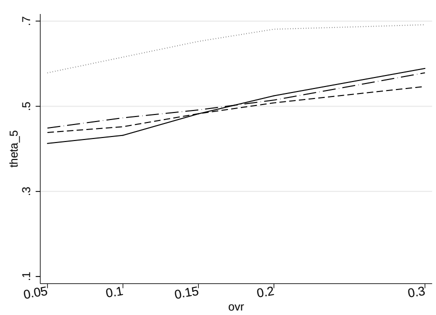
(b) $\rho^2 = 0.3$



(c) $\rho^3 = 0.5$



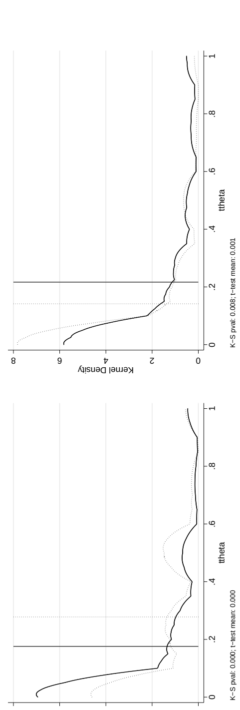
(d) $\rho^4 = 0.7$



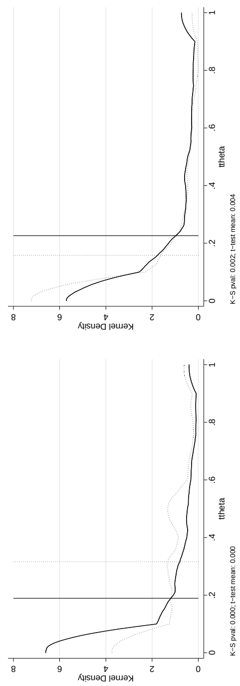
(e) $\rho^5 = 0.9$

Figure 3.8: Average Percentage of Income Invested by Option

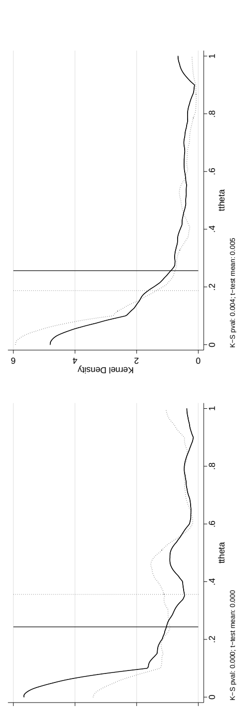
Source: solid $(\overline{w}, L_1)$, dotted $(\underline{w}, L_1)$, dashed $(\overline{w}, L_0)$,dashed–dotted $(\underline{w}, L_0)$



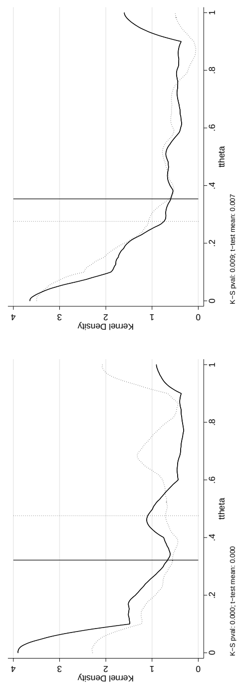
(a) $\rho^1 = 0.1$. $\underline{w}$ (left panel) $\overline{w}$ (right panel)



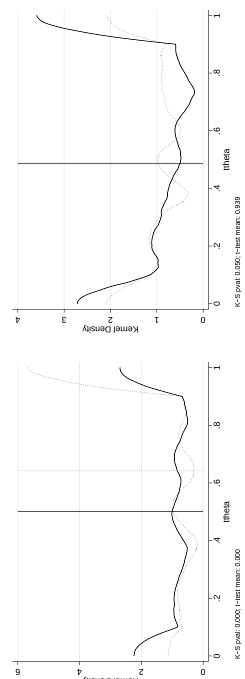
(b) $\rho^2 = 0.3$. $\underline{w}$ (left panel) $\overline{w}$ (right panel)



(c) $\rho^1 = 0.5$. $\underline{w}$ (left panel) $\overline{w}$ (right panel)



(d) $\rho^1 = 0.7$. $\underline{w}$ (left panel) $\overline{w}$ (right panel)



(e) $\rho^1 = 0.9$. $\underline{w}$ (left panel) $\overline{w}$ (right panel)

Figure 3.9: Kernel Bounded Density of Income Invested
Source: L_0 (solid), L_1 (dotted). Vertical line represents mean

3.7.3 Experimental instructions

Instructions for Monitors

1. In the Excel file, set Save As, and give the name of the file ID of the participant.
2. The sheet "ID" is only for information and can not be seen by the participant, ensure you have this part ready before the participant arrives.
3. There are two different treatments and should be implemented accordingly with the different instructions at the beginning of the game. The code used in the excel file will be Income Treatment (IT): If (1) the wealth will be $ 20,000, if (2) is $ 40,000, if (3) is $ 20,000 or $ 40,000 as the result issued by the coin toss (3). For (3) let the participant choose which side of the coin will associated to each amount. However the coin should fall to the floor (this is to ensure participants cannot influence the coin toss) After the allocation of income, you need to put on the table the player's initial income and tell him that that money is already of his own.
4. At the beginning of each round you must follow the following wording: "We are in the XX round, the probability of business failure is XX%, which is the same as taking a red ball from the bag that we will fill with 10 balls, among which XX/10 are red balls and (100-XX)/10 are green, equivalently the probability of being successful is 1-XX%, which is the same as taking a green ball from the bag".
5. In Round XX, in Step 1 of each alternative scenario start saying "How much would you be willing to invest in the business in Scenario XX in which you could lost 90% of your investment if it fails or get a return YY% if successful?"
6. Once the individual has responded continue saying, in the most neutral possible way to avoid inducing an answer, "You just told me you would be willing to invest $ XX, equal to XX% of your current income, this means you will lose XX (which is the same as keeping just XX from the investment amount side) if the business fails or win XX (which is the same as keeping XX from the invested amount) if the business is successful. Do you want to keep or change the amount invested in this scenario?" If participant decides to change the amount invested, repeat 5 and 6.
7. Once step 1 is completed for that scenario you move on to step 2. You should tell the participant "Now we give you the offer of an insurance against your investment. It gives you back the expected value of your investment minus a fixed value $X. What do you prefer between: Option A, to stay with you investment or Option B, the insurance that would return you a total amount of $ Y." Once the participant has answered then ask him, again in a neutral mood, "Do you want to keep or change your choice between Option A and Option B in this scenario?". Repeat 5, 6 and 7 to complete the five scenarios. Always remember to code Step 1 (amount to be invested in the business) and then Step 2 (investment vs fair insurance) for all the five scenarios.

IMPORTANT: If at any time the player asks you to return to review any of the five decisions within a same round, you can do so only if it is in the same round, however there is no option to change previous decisions.

8. upon completion of the five alternative scenarios, go to Table 5 which summarizes the scenarios and show it to the participant.
9. Now give the participant 5-side die. Describe to the participant the randomly chosen scenario. That is, if you get from rolling the die the stage 5, show him that stage and summarise to him his decision: “you decided to invest $ x dollars in business and chose the option (A / B).

10. If in the relevant scenario the participant chose Option A, take the black bag and include green and red balls accordingly with the stated probability of success. Do it transparent to the participant. Stir it, and let the player draw a ball, make sure it does not look inside the bag.
11. Summarize to the participant the result.
12. Finish the round t reminding him/her that his income at the end of this round will be his/her initial wealth for round t+1.
13. Save the file at the end of each round to protect against data loss.

To bear in mind:
• Remember to stay as neutral as possible. Do not induce responses. Call the experiment, activity or game.
• If you are asked about the activity or research project aim, tell them that we want to characterize the decisions of a group of people. But tell them that at the end of the activity they will be invited to express their comments and that one member of the research group will be happy to answer their questions.
• After they have confirmed a decision it CANNOT be changed afterwards. If they say no longer maintain it may change at any time. Not when they are taking another decision and want to see that or when playing step 2.
• Avoid telling the business "uncertain business" or "lottery". Just call it "business" or “alternative”.

Instruction for participants Stage 2. (I)

This activity will be implemented in a computer with the assistance of a monitor. You begin the activity with an initial income in real money that will be assigned by the monitor. You receive this wealth only after the whole the activity has finished, after we have add your gains and subtract your losses. However, **the initial income is already yours since the beginning because you have reached to this point of the activity.**

The activity consists of several rounds; in each round you must take five (5) decisions. At the end of each round you would obtain gain or losses and those will depend on your choices, the outcome at the end of the round will be add or subtracted from the income you had at the beginning of that round. The number of rounds will be determined by the computer, and neither the monitor nor you will know this information in advance.

Let's see how you can make choices at each round

After the monitor reports your initial income, you must **carry out five (5) decisions**, one for each of five Alternatives (Alternative 1, Alternative 2... until Alternative 5). To determine the outcome at the end of the round, only one of those decisions will be taken into account. However, before taking decisions **YOU WILL NOT KNOW** which Alternative, form the five Alternatives, is the one to be taken into account.

The alternative taken into account will be chosen at random throwing a 10-sided die ranging from 0 to 9. If you get a value into 1 or 2 will be taken into account the Alternative 1, if you get a value into 3 or 4 will be taken into account the Alternative 2 and so on, thus if the die throws in the value 9 or 0 will be taken into account is Alternative 5 (face to 0 is the value 10).

Each of the five (5) decisions consists of two (2) steps. Let's see:

Step 1.

You must determine how much of the starting income you want to invest if given each business alternative. You are free to choose not to invest, to invest all your wealth or invest only part of it the business. Observe that the maximum amount that you can invest in each alternative is the starting income from that round. It is important to know that the money you choose not to invest will be there for next round without any modification. In other words, you do not lose the money you decided not to invest. Before we proceed further let us clarify what is a business alternative.

Any business alternative has two possible outcomes: i) **Success:** If it is successful you will get back the amount of money invested plus an additional return. The return will be between 5% and 30% depending on the alternative and you will know it in advance. Hence, if the business is successful, for every 100 pesos invested, you will get back those 100 pesos plus something between 5 and 30 pesos), or ii) **Failure**: in which case, you will lose 90% of the amount of money you decided to invest. So if you invested 100 pesos then you will lose 90 pesos and will keep 10 pesos.

Step 2

You must choose between two options: **A to keep** the investment on the business Alternative as an investment you took in step 1, or **B** change the business by acquiring insurance, which give your invested money back minus a the price of the insurance.

The five business alternatives in which you must make the decision differ on the rate of return you would get if the business is successful. However, the amount you would lose, if the business fails, will always be the same, equivalent to losing 90% of the amount invested. Thus:

- If Alternative 1 is chosen, after the die is thrown, the business will have a return of 5% (for every 100 pesos spent you get additional 5 pesos)
- If Alternative 2 is chosen, after the die is thrown, the business will have a return of 10% for every 100 pesos spent you gets additional 10 pesos)
- If Alternative 3 is chosen, after the die is thrown, the business will have a return of 15% for every 100 pesos spent you gets additional 15 pesos)
- If Alternative 4 is chosen, after the die is thrown, the business will have a return of 120% for every 100 pesos spent you gets additional 20 pesos)
- If Alternative 5 is chosen, after the die is thrown, the business will have a return of 30% for every 100 pesos spent you gets additional 30 pesos)

It is important to note that in the same round, investing in an alternative does not affect the income you for investing in the other alternatives within the same round. They are independent alternatives.

Likewise, at the beginning of each round you will be informed of the chance or probability that the business is successful (represented by the green balls.) or failure (represented by red balls.). These chances will be different between rounds.

Now let us show you an example of the decision making process for the first round

Example 1

i. The monitor will give you an initial income. Suppose it is $ 2,000 pesos.

ii. Since we are in Round 1, you will see on the computer the Characteristics of Business. At the top of the screen you will find the following:

You will see the staring income for this round, the chance, for this round, of the business failing (in this case equal to nine red balls) and the chance of it succeeding (in this case a green ball).

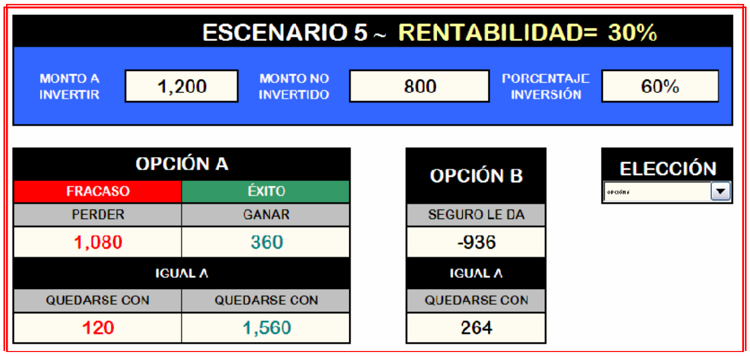
Knowing this information, you are know ask to determine how much of $ 2,000 you want to invest in each of the five alternatives, and for each of them you have decided whether you prefer staying with the investment or you want to change for the insurance and pay the price.

iii. Once you have taken the five (5) decisions for each alternative, we will throw the 10-sides die to determine which of the alternatives will be taken into account to determine the outcome. Suppose that the die get on 9, thus the considered alternative is the fifth one.

iv. We go to alternative 5 an check your decision made. Suppose that in this scenario you decided to invest $ 1,200pesos (and therefore left without investing $ 800 pesos) in step 1. So you would have the following

Accordingly, this investment would get $ 360 pesos if the business is successful equivalent to keeping $ 1,560 ($ 1,200 invested plus $ 360 profit), and would lose $ 1,080 (equivalent to keep $120 of that $ 1,200 invested). The probability of each outcome is given by the number of green and red balls respectively.

Suppose further that in step 2 you chose to keep the investment (Option A) rather than swift to the insurance (Option B), thus you will see the following



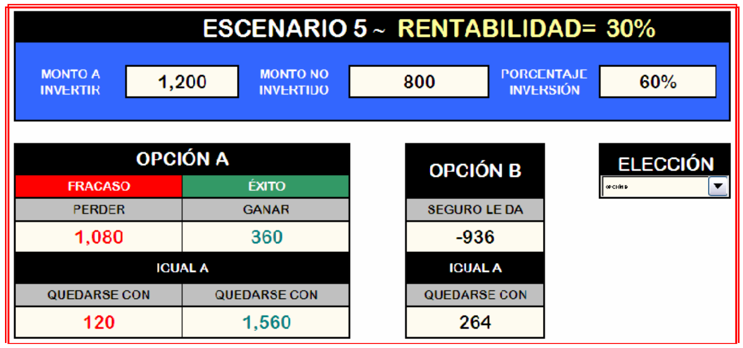
Therefore, it is necessary to determine what would happen with the business in that round

v. Inside a black bag we put nine red balls and one green ball (because that is equivalent to the probability of failure and success respectively). You will draw one of the balls.

vi. Suppose that the ball was green. Then we can get compute your income at the end of Round 1. Your initial wealth was $ 2,000 pesos. The die determined that the Alternative 5 would be taken into account. Your decision in Step 1 was to invest $1,200 pesos, and in the step 2 you decided to choose the investment rather than the insurance. The ball you picked determined that the investment was successful. Then, your investment of $ 1,200 pesos became $1,560 pesos. Thus, your wealth at the end of Round 1 is $ 2,360. This represents the sum of $ 800 not invested in the business plus $ 1,560 you won with this business.

Suppose now the ball drawn is red instead. Then, the result of the investment is failure. Like we have already explained, , you lose 90% of the amount invested equivalent to $1,080. Therefore, will stay with the $ 800 (which is not invested) plus $ 120 representing 10% of which are kept its investment. Consequently, your wealth at the end of the round 1 is $ 920.

Suppose finally a third possibility. Everything remains the same but in step 2 instead you chose option B (insurance). In this case, the Step 2 will look like



In this case, it is not necessary to run the business according to the probabilities (no need to darw one ball form the black bag). This is so because you decided to go for the insurance (Option B). In this case, your final income in Round 1 will be the sum of the amount investment $ 1,200 pesos minus the price of the insurance $936 equivalent to $ 264 ($ 1,200 - $ 936 = $ 264) plus $800 pesos not invested in the business, leaving a total amount of $ 1.064.

One Round 1 has ended you will be informed of your new income. The following rounds remain exactly as round 1, but with one difference: the initial income of each of these rounds will not be assigned by the monitor, instead it will be equal to the final income from the previous round.

Therefore, Initial income at round 1 will be allocated by the monitor, Initial income at round 2 will be the initial income of round 1 plus gains or minus losses form the investment decision at round 1, which is the same as the final income at round 1. Initial income at round 3 is the final income of round 2, and so on and so forth. Additionally, you will be informed about the probability of success and failure at each round.

You need to know that all rounds will be played consecutively and immediately, that is, once you finish Round 1, Round 2 will start and you will make decisions corresponding to the alternatives of that round, after Round 2 finishes then you will move to Round 3, and so on and so forth until the end of the rounds. The income you earned at the end of the last round will be the gain in this activity. Remember that payments will be made in cash (Colombian pesos). We guarantee that we are in full capacity to pay what you earn in the activity. We must say that during the rounds there will be no delivery of money but only at the end of all the rounds.

If you have any questions this is the time to make.

Now you will be accompanied with one monitor that will lead you across the rounds. You will play one practice round with a hypothetical income. After you have completed that round you will begin playing with real income that the monitor will assign to you.

Instruction for participants Stage 2. (II, Reference point treatment fixed)

Reference Point Treatment = Fixed

(Once subjects have completed the practice round)

You will start this stage with an income equal to $ 40,000 (20,000). This amount of money is already yours given you have reached to this point in the activity. Your decisions must be taken with reference to this income. Depending on your decisions and the outcomes you get, we will add gains or losses to determine the final income.

Reference Point Treatment = Random

(Once subjects have completed the practice round)

You could start this stage with an income equal to $ 40,000 or $ 20,000. The actual amount will depend on a coin toss. You must decide which side of the coin is associated with each income level, once assigned you will toss the coin.

Please assign an income level to each side of the coin

(Once subjects have assigned income levels to sides of the coin)

Please, toss the coin

(Once subjects have tossed the coin)

According to the coin flip you have loss the chance of starting with an income equal to $ 40,000 (20,000) instead you will start this stage with an income equal to $ 40,000 (20,000). This amount of money is already yours given you have reached to this point in the activity. Your decisions must be taken with reference to this income. Depending on your decisions and the outcomes you get, we will add gains or losses to determine the final income. is amount of money is already yours given you have reached to this point in the activity. Your decisions must be taken with reference to this income. Depending on the outcomes, we will add gains or losses obtained across rounds.

Instructions for participants Stage 1

In this activity, your earnings depend on your decisions. There is no wrong or correct answer. There will be three (3) set of situations. The first set (denoted as **Set 1**) consists of 14 situations; the second set (**Set 2**) consists of 14 situations and the **Set 3** of 7 situations. Consequently, there are a total of 35 situations (please see the record sheet). In every Situation we offer two choices: Option A and Option B. **In each Situation we ask that you choose either Option A or Option B according to what you prefer.**

After you have completed your decisions, we will put into a bag 35 black chips numbered and you will select one at random to select a Situation from the 35 situations which you answered. The Situation selected will be played for real money.

For example, if the number selected is 15, we will apply the Situation 15 with real money. Once the Situation is drawn we take into account your Option (A or B) chosen. Then we will include in a black bag 10 ping-pongs between green (●) and red (●). Where the number of green balls (●) represents the chance that you obtain the greatest amount, and the red balls (●) that you of the lower amount. Once you draw a ball we will give you the amount of money.

Now, let us provide an example of the **Set 1**.

Example 1

This example is for the Set 1, please look at your record sheet. (Situations from 1 to 14)

Note that there are two Options, A and B for each Situation. In a black bag there are 10 balls between green ● and red ●. You must choose between A or B in each Situation. Suppose that the number drawn at random is number 1, thus we take your decision made at Situation 1.

Situation	Option A	Option B
1	$ 30.000 if ❶❷❸ $ 7.500 if ❶❷❸❹❺❻❼	$ 47.000 if ❶ $ 3.750 if ❶❷❸❹❺❻❼❽❾

If you chose Option A, the bag will contain three (3) green balls ● and seven (7) red balls ●. On the other hand, if elected Option B the bag containing one (1) green ball ● and nine (9) red balls ●. Afterwards, you will draw a ball from the bag at random.

If the green ball ● is drawn, those who chose Option A would receive $ 30,000 and those who chose Option B will receive $ 47,000.

If the red ball ● is drawn, those who chose Option A will receive $ 7,500 and those who chose Option B will receive $ 3,750

Example 2

Suppose now a situation where you choose Option A from Situation 1 to 6 and Option B from Situation 7 to 14. Then, you must complete your record sheet as follows:
Sets 1

Situation	Option A	Option B
1	$ 30.000 si ❶❷❸ $ 7.500 si ❶❷❸❹❺❻❼	$ 47.000 si ❶ $ 3.750 si ❶❷❸❹❺❻❼❽❾
2	$ 30.000 si ❶❷❸ $ 7.500 si ❶❷❸❹❺❻❼	$ 52.000 si ❶ $ 3.750 si ❶❷❸❹❺❻❼❽❾
3	$ 30.000 si ❶❷❸ $ 7.500 si ❶❷❸❹❺❻❼	$ 59.000 si ❶ $ 3.750 si ❶❷❸❹❺❻❼❽❾
4	$ 30.000 si ❶❷❸ $ 7.500 si ❶❷❸❹❺❻❼	$ 67.000 si ❶ $ 3.750 si ❶❷❸❹❺❻❼❽❾
5	$ 30.000 si ❶❷❸ $ 7.500 si ❶❷❸❹❺❻❼	$ 78.000 si ❶ $ 3.750 si ❶❷❸❹❺❻❼❽❾
6	$ 30.000 si ❶❷❸ $ 7.500 si ❶❷❸❹❺❻❼	$ 91.000 si ❶ $ 3.750 si ❶❷❸❹❺❻❼❽❾
7	$ 30.000 si ❶❷❸ $ 7.500 si ❶❷❸❹❺❻❼	$ 109.000 si ❶ $ 3.750 si ❶❷❸❹❺❻❼❽❾
8	$ 30.000 si ❶❷❸ $ 7.500 si ❶❷❸❹❺❻❼	$ 125.000 si ❶ $ 3.750 si ❶❷❸❹❺❻❼❽❾
9	$ 30.000 si ❶❷❸ $ 7.500 si ❶❷❸❹❺❻❼	$ 149.000 si ❶ $ 3.750 si ❶❷❸❹❺❻❼❽❾
10	$ 30.000 si ❶❷❸ $ 7.500 si ❶❷❸❹❺❻❼	$ 184.000 si ❶ $ 3.750 si ❶❷❸❹❺❻❼❽❾
11	$ 30.000 si ❶❷❸ $ 7.500 si ❶❷❸❹❺❻❼	$ 238.000 si ❶ $ 3.750 si ❶❷❸❹❺❻❼❽❾
12	$ 30.000 si ❶❷❸ $ 7.500 si ❶❷❸❹❺❻❼	$ 330.000 si ❶ $ 3.750 si ❶❷❸❹❺❻❼❽❾
13	$ 30.000 si ❶❷❸ $ 7.500 si ❶❷❸❹❺❻❼	$ 509.000 si ❶ $ 3.750 si ❶❷❸❹❺❻❼❽❾
14	$ 30.000 si ❶❷❸ $ 7.500 si ❶❷❸❹❺❻❼	$ 923.000 si ❶ $ 3.750 si ❶❷❸❹❺❻❼❽❾

Answer to Set 1

I chose Option A from situation **1** **to** 6

I chose Option B from situation 7 **till** **14**

Example 3
Now le tus suppose that you choos Option A for all 14 situations, then you must fill out your record sheet as follows:

Situación	Option A	Option B
1	$ 30.000 si ❶❷❸ $ 7.500 si ❶❷❸❹❺❻❼	$ 47.000 si ❶ $ 3.750 si ❶❷❸❹❺❻❼❽❾
2	$ 30.000 si ❶❷❸ $ 7.500 si ❶❷❸❹❺❻❼	$ 52.000 si ❶ $ 3.750 si ❶❷❸❹❺❻❼❽❾
3	$ 30.000 si ❶❷❸ $ 7.500 si ❶❷❸❹❺❻❼	$ 59.000 si ❶ $ 3.750 si ❶❷❸❹❺❻❼❽❾
4	$ 30.000 si ❶❷❸ $ 7.500 si ❶❷❸❹❺❻❼	$ 67.000 si ❶ $ 3.750 si ❶❷❸❹❺❻❼❽❾
5	$ 30.000 si ❶❷❸ $ 7.500 si ❶❷❸❹❺❻❼	$ 78.000 si ❶ $ 3.750 si ❶❷❸❹❺❻❼❽❾
6	$ 30.000 si ❶❷❸ $ 7.500 si ❶❷❸❹❺❻❼	$ 91.000 si ❶ $ 3.750 si ❶❷❸❹❺❻❼❽❾
7	$ 30.000 si ❶❷❸ $ 7.500 si ❶❷❸❹❺❻❼	$ 109.000 si ❶ $ 3.750 si ❶❷❸❹❺❻❼❽❾
8	$ 30.000 si ❶❷❸ $ 7.500 si ❶❷❸❹❺❻❼	$ 125.000 si ❶ $ 3.750 si ❶❷❸❹❺❻❼❽❾
9	$ 30.000 si ❶❷❸ $ 7.500 si ❶❷❸❹❺❻❼	$ 149.000 si ❶ $ 3.750 si ❶❷❸❹❺❻❼❽❾
10	$ 30.000 si ❶❷❸ $ 7.500 si ❶❷❸❹❺❻❼	$ 184.000 si ❶ $ 3.750 si ❶❷❸❹❺❻❼❽❾
11	$ 30.000 si ❶❷❸ $ 7.500 si ❶❷❸❹❺❻❼	$ 238.000 si ❶ $ 3.750 si ❶❷❸❹❺❻❼❽❾
12	$ 30.000 si ❶❷❸ $ 7.500 si ❶❷❸❹❺❻❼	$ 330.000 si ❶ $ 3.750 si ❶❷❸❹❺❻❼❽❾
13	$ 30.000 si ❶❷❸ $ 7.500 si ❶❷❸❹❺❻❼	$ 509.000 si ❶ $ 3.750 si ❶❷❸❹❺❻❼❽❾
14	$ 30.000 si ❶❷❸ $ 7.500 si ❶❷❸❹❺❻❼	$ 923.000 si ❶ $ 3.750 si ❶❷❸❹❺❻❼❽❾

Answer to Set 1

I chose Option A from situation 1 to $\boxed{14}$

~~**I chose Option B from situation** $\boxed{}$ **till 14**~~

Now lets see an experiment with **Set 2**. **Please refer to the record sheet. (Situations 15 until 28)**

Example 4

Lets assume that you chose Option B for all situations, therefore you must fill out the record sheet as follows:

Set 2

Situación	Option A	Option B
15	$ 20.000 si ❶❷❸❹❺❻❼❽❾ $ 15.000 si ❶	$ 26.000 si ❶❷❸❹❺❻❼ $ 3.750 si ❶❷❸
16	$ 20.000 si ❶❷❸❹❺❻❼❽❾ $ 15.000 si ❶	$ 27.000 si ❶❷❸❹❺❻❼ $ 3.750 si ❶❷❸
17	$ 20.000 si ❶❷❸❹❺❻❼❽❾ $ 15.000 si ❶	$ 28.100 si ❶❷❸❹❺❻❼ $ 3.750 si ❶❷❸
18	$ 20.000 si ❶❷❸❹❺❻❼❽❾ $ 15.000 si ❶	$ 29.000 si ❶❷❸❹❺❻❼ $ 3.750 si ❶❷❸
19	$ 20.000 si ❶❷❸❹❺❻❼❽❾ $ 15.000 si ❶	$ 30.000 si ❶❷❸❹❺❻❼ $ 3.750 si ❶❷❸
20	$ 20.000 si ❶❷❸❹❺❻❼❽❾ $ 15.000 si ❶	$ 31.000 si ❶❷❸❹❺❻❼ $ 3.750 si ❶❷❸
21	$ 20.000 si ❶❷❸❹❺❻❼❽❾ $ 15.000 si ❶	$ 33.000 si ❶❷❸❹❺❻❼ $ 3.750 si ❶❷❸
22	$ 20.000 si ❶❷❸❹❺❻❼❽❾ $ 15.000 si ❶	$ 34.000 si ❶❷❸❹❺❻❼ $ 3.750 si ❶❷❸
23	$ 20.000 si ❶❷❸❹❺❻❼❽❾ $ 15.000 si ❶	$ 39.000 si ❶❷❸❹❺❻❼ $ 3.750 si ❶❷❸
24	$ 20.000 si ❶❷❸❹❺❻❼❽❾ $ 15.000 si ❶	$ 45.000 si ❶❷❸❹❺❻❼ $ 3.750 si ❶❷❸
25	$ 20.000 si ❶❷❸❹❺❻❼❽❾ $ 15.000 si ❶	$ 51.000 si ❶❷❸❹❺❻❼ $ 3.750 si ❶❷❸
26	$ 20.000 si ❶❷❸❹❺❻❼❽❾ $ 15.000 si ❶	$ 57.000 si ❶❷❸❹❺❻❼ $ 3.750 si ❶❷❸
27	$ 20.000 si ❶❷❸❹❺❻❼❽❾ $ 15.000 si ❶	$ 62.000 si ❶❷❸❹❺❻❼ $ 3.750 si ❶❷❸
28	$ 20.000 si ❶❷❸❹❺❻❼❽❾ $ 15.000 si ❶	$ 67.000 si ❶❷❸❹❺❻❼ $ 3.750 si ❶❷❸

Answer to Set 2

~~**I chose Option A from situation 1 to**~~ ☐

I chose Option B from situation 1 **till 14**

Now let us consider **Set 3**. Note **Set 3** situations include losses. **Please refer to the record sheet. (Situations 29 until 35)**

Set 3

Situación	Option A	Option B
29	Gana $ 25.000 si ❶❷❸❹❺ Pierde $ 4.000 si ❶❷❸❹❺	Gana $ 30.000 si ❶❷❸❹❺ Pierde $ 21.000 si ❶❷❸❹❺
30	Gana $ 4.000 si ❶❷❸❹❺ Pierde $ 4.000 si ❶❷❸❹❺	Gana $ 30.000 si ❶❷❸❹❺ Pierde $ 21.000 si ❶❷❸❹❺
31	Gana $ 1.000 si ❶❷❸❹❺ Pierde $ 4.000 si ❶❷❸❹❺	Gana $ 30.000 si ❶❷❸❹❺ Pierde $ 21.000 si ❶❷❸❹❺
32	Gana $ 1.000 si ❶❷❸❹❺ Pierde $ 4.000 si ❶❷❸❹❺	Gana $ 30.000 si ❶❷❸❹❺ Pierde $ 16.000 si ❶❷❸❹❺
33	Gana $ 1.000 si ❶❷❸❹❺ Pierde $ 8.000 si ❶❷❸❹❺	Gana $ 30.000 si ❶❷❸❹❺ Pierde $ 16.000 si ❶❷❸❹❺
34	Gana $ 1.000 si ❶❷❸❹❺ Pierde $ 8.000 si ❶❷❸❹❺	Gana $ 30.000 si ❶❷❸❹❺ Pierde $ 14.000 si ❶❷❸❹❺
35	Gana $ 1.000 si ❶❷❸❹❺ Pierde $ 8.000 si ❶❷❸❹❺	Gana $ 30.000 si ❶❷❸❹❺ Pierde $ 11.000 si ❶❷❸❹❺

If you get losses then they will be substracted from the final amount of money you get from both Stages in the overall Activity (Stage 1 and Stage 2).

If you have any question this is the moment to raise it.

It is important to bear in mind that you must first fill out the record sheet according to your preferences from Stage 1. Afterwards, you will come back to the main room where we will deliver the instructions for the Stage 2. Once we finish Reading them each of you will be assigned a Monitor to help you out in the implementation of your decisión from that Stage 2. After you have finished Stage 2 each Monitor will have all your record sheets from Stage 1 and he will implement the corresponding payments. We guarantee that we are able to pay any amount that you have earned during the experiment.

With Set 1 and Set 2 we recover: Measure of decreasing sensitivity (γ) and weighting of probabilities (π)

To do so we follow the same procedure described in Tanaka et al. (2010). We find the associated $\alpha_i, \lambda_i, \gamma_i$ parameters that rationalises the $V_i(L_{A,a}) - V_i(L_{B,a}) = 0$ where $L_{x,a}$ is the lottery associated to Option X in row a, and $V_i(L_X) = (\pi_i(\rho)\mu_i(x \mid b)) \mid_{L_X}$ were

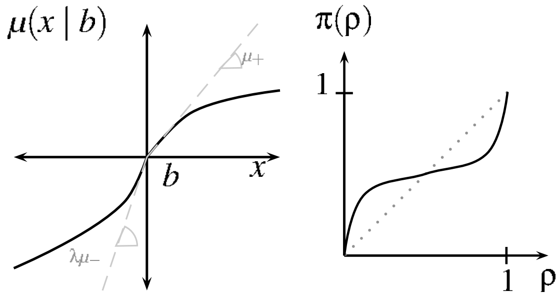
$$\mu_i(x \mid b) = \begin{cases} \frac{x^{1-\gamma_i}}{1-\gamma_i} & \text{if } x \geq 0, \\ \frac{-\lambda_i(-x)^{1-\gamma_i}}{1-\gamma_i} & \text{if } x < 0 \end{cases}$$

weighted by $\pi_i(\rho) = \exp^{[-(-\log(\rho))_i^{\alpha}]}$. Notice that set 1 and set 2 gives us an interval of values for which, at a given switching point, individual is consistent with a duple (γ, α). Set 3 is used to recover parameter λ_i.

Set 1 switching points in Table 3.6

Set 2 switching points in Table 3.7

With Set 3 we recover loss aversion (λ).



(a) Gain–loss utility $\mu(\cdot)$ (b) Probability weighting: $\pi(\rho)$

Table 3.6: Set 1. switching points α y γ

$\sigma\backslash\alpha$	**0.4**	**0.5**	**0.6**	**0.7**	**0.8**	**0.9**	**1**
0.2	9	10	11	12	13	14	NC*
0.3	8	9	10	11	12	13	14
0.4	7	8	9	10	11	12	13
0.5	6	7	8	9	10	11	12
0.6	5	6	7	8	9	10	11
0.7	4	5	6	7	8	9	10
0.8	3	4	5	6	7	8	9
0.9	2	3	4	5	6	7	8
1	1	2	3	4	5	6	7

*never changes

Table 3.7: Set 2. Switching point for α y γ

$\sigma \backslash \alpha$	**0.4**	**0.5**	**0.6**	**0.7**	**0.8**	**0.9**	**1**
0.2	NC*	14	13	12	11	10	9
0.3	14	13	12	11	10	9	8
0.4	13	12	11	10	9	8	7
0.5	12	11	10	9	8	7	6
0.6	11	10	9	8	7	6	5
0.7	10	9	8	7	6	5	4
0.8	9	8	7	6	5	4	3
0.9	8	7	6	5	4	3	2
1	7	6	5	4	3	2	1

*Never changes

Chapter 4

Interdependent Value Auctions with Insider Information

4.1 Introduction

Much of the auction literature assumes that bidders hold rather equally informative information about the value of the auctioned object, though they are assumed to know about it privately.[1] However, there are real-world auctions in which bidders are better viewed as *asymmetrically* informed in terms of how precisely they know about different aspects of the object's value. For instance, in art auctions, buyers with professional knowledge tend to appraise more accurately the potential value of an object than non-professional buyers (Ashenfelter & Graddy 2003). In takeover auctions, buyers with the existing shares of a target firm may have an access to inside information unavailable to competitors.[2] In auctions for gas and oil leases (so-called OCS auctions), firms owning neighboring tracts are better informed about the value of a lease, such as oil reserves, than non-neighboring firms (Hendricks, Porter & Wilson 1994). In these examples, some bidders have far better access to value information than others, which suggests the asymmetry of information access among bidders. This motivates us to explore the implications of this asymmetry on auction outcomes.

Further elaborating on Kim (2003) we develop a model of interdependent value

[0]The chapter has benefited from comments by seminar participants at UCL Theory Seminar and 2013 ESA Zurich Meeting. We thank Brian Wallace for writing the experimental program and helping us run the experiment.

[1]See, for instance, Myerson (1981) and Riley & Samuelson (1981) for the environment of private value auctions, and Milgrom & Weber (1982*a*) for auctions with affiliated values.

[2]Sometimes, a current management team of the target firm participates in the bidding competition, which is a practice known as management buyouts or MBO. Shleifer & Vishny (1987) argue that the managers' special information about their company is one reason for MBO.

auctions with *asymmetric information structure* in two standard auction formats – the second-price (sealed-bid) auction and the English auction. The model assumes a flexible structure of valuation, encompassing both private component specific to individual bidder and common component applied to all bidders. We depart from the standard structure by assuming two types of bidders: *insiders* who are perfectly informed of their value, and *outsiders* who are only informed of the private value component and thus imperfectly informed about their value. Our theoretical analysis is general in that we allow for any arbitrary number of insiders and outsiders in the equilibrium analysis. This feature enables us to understand the revenue implication of the insiders' presence by examining how the seller's revenue changes as the number of insiders varies.

When only one insider and one outsider exists, the second-price auction and the English auction, which are equivalent, achieve the efficient allocation. With more than two bidders and at least one insider, however, the equilibrium of the second-price auction is inefficient. Efficiency loss is unavoidable in the second-price auction because a sealed-bid format offers no way of eliminating the mismatch between insiders' and outsiders' bidding strategies, which arises from the fact that the former depends on both private and common components of the values while the latter depends only the private component.

The English auction in contrast provides opportunities for outsiders to learn about others'—both insiders and outsiders'—private information through the history of prices at which they drop out. In an environment with symmetric information structure (that is, there is no insider), each bidder's private signal is precisely revealed in equilibrium because his equilibrium drop-out price is monotonically related to his private signal. Yet the inference problem can be more complicated in the presence of insiders, who employ the (weakly) dominant strategy of dropping out at their values that reflect both private and common components. We characterize an equilibrium in English auction where outsiders overcome this inference problem to get the object allocated efficiently.

We do so by extending the equilibrium construction in Milgrom & Weber (1982*b*) and Krishna (2003) to our setup, which involves finding a system of equations that gives *break-even signals* at any given price in the presence of insiders. The outsider's equilibrium strategy is then characterized by the cutoff strategy in which he drops out at such a price that the break-even signal becomes equal to his signal. The key to the equilibrium construction is to guarantee that each outsider's break-even signal as a function of the price is strictly monotone, which holds true in our setup though it requires a non-trivial analysis. The outsiders' drop-out strategy, based on the monotonic break-even signals, then implies that outsiders drop out before the price reach their own value and drop out in order of their values. An immediate consequence of this strategy and insiders' dominant strategy is that in the case an insider has the highest value, he becomes a winner. In the case where an outsider has the highest value, the allocation is also efficient because all other outsiders drop out earlier, revealing their precise signals, and thus the informational gap between the highest-value outsider and any remaining insider is reduced sufficiently to attain the efficient allocation. Our results highlight the efficiency-achieving role of the dynamic nature of the English auction in an interdependent value setup with asymmetric structure of information.

We also explore the revenue implications of the asymmetric information structure in the English auction. To investigate this, we consider two English auctions that differ only by one bidder who switches from an outsider in one auction to an insider in the other auction. With the equilibrium of the English auction we construct, we show that for any signal profile, the latter auction with the switched insider yields a (weakly) higher revenue than the former one. This result is based on two effects of turning an outsider into an insider on bidders' bidding behaviour. First, the switched insider drops out at a higher price than when he is an insider. Second, the higher drop-out price of the switched insider in turn causes active outsiders to drop out at higher prices. This revenue prediction is reminiscent of the linkage principle of Milgrom & Weber (1982*b*) in that the switch of an outsider into insider increases information available to bidders. However, there are important differences between our revenue result and Milgrom and Weber's prediction. Unlike the case of Milgrom & Weber (1982*b*) where the extra information available to bidders is made public, only one bidder entertains better information in our setup. Also, our revenue prediction holds *ex-post*—that is, for every realization of the signal profile— and thus does not depend on the assumption that signals are affiliated.

While our theory offers a nice benchmark to compare the efficiency and revenues of two standard auction formats in the environment with asymmetric information structure, its predictions are based on the nontrivial inference process in the English auction. Whether individual bidders can rationally process the information revealed through the auctions and the theoretical predictions are thus valid is ultimately an empirical question. To test the validity of the theory, we design a simple experiment by varying the auction format—between English and second-price—and the composition of insiders and outsiders. Specifically, we employ three-bidder auctions of each format with either three outsiders, or two outsiders and one insider, or one outsider and two insiders. Each combination of an auction format and an insider-outsider composition serves as a single treatment. In order to make outsiders' inference problem as transparent as possible, we let the computer play the role of insider who follows the dominant strategy of dropping out at its own value This was public information to all human subjects.

Our experiment presents several findings. First, the English auction achieves a higher level of efficiency than the second-price auction when at least one insider is present. And there is no significant difference of efficiency between the two auctions when there is no insider. This finding on efficiency is consistent with our theory. Second, average revenues tend to deviate upward from the equilibrium benchmark, in particular, in both auction formats with no insider. Despite this, the increase in the number of insiders has a positive impact on revenues in the English auction, as the theory predicts. We also found a similar pattern of the increase of revenues in the second-price auction with respect to the number of insiders. Third, there is evidence of naive bidding in both auction formats, that is, winners' curse behaviour. Quite intriguingly, the degree of naive bidding in the data declines in the number of insiders in both auction formats. In particular, this pattern is statistically significant and behaviourally large in the English auction. We conjecture that the presence of insiders—who have informational advantage—makes the outsider be more careful in bidding and creates a behavioural incentive for the outsider to hedge against the informational disadvantage. This can work toward the correction of naive bidding and thus of the winner's curse in our setup.

Literature: This chapter contributes to the literature of auctions with asymmetric information structure. Engelbrecht-Wiggans, Milgrom & Weber (1983) study a first-price common-value auction in which a single 'insider' has proprietary information about the common value of the object, while other bidders have public information. Hendricks et al. (1994) and Hendricks & Porter (1988) extend the analysis of Engelbrecht-Wiggans et al. (1983) and study oil and gas drainage lease auctions. Also, in first-price common-value auctions, Campbell & Levin (2000) and Kim (2008) examine theoretically the effects of an insider on revenues and Kagel & Levin (1999) study experimentally the effects of an insider on revenues and bidding behaviour. Our chapter differs in the above papers in some important manners. We study the implications of insider information on efficiency as well as revenues in interdependent value auctions, whereas the literature focuses on the revenue implications due to the pure common-value assumption. We study the English auction and the second-price auction rather than the first-price auction. And we provide a general equilibrium analysis for any arbitrary number of insiders and generate the revenue ramifications of introducing an extra insider, whereas the literature allows only a single insider.

Our model brings new insights to the literature concerning the allocative efficiency in the interdependent value environment. Dasgupta & Maskin (2000), and Perry & Reny (2002) design some original mechanisms that implement the efficient allocation under a fairly weak condition (single crossing property). Krishna (2003) studies the

English auction with interdependent values and asymmetric bidders in valuation and adapts the equilibrium characterization of Milgrom & Weber (1982*a*). He provides sufficient conditions for the existence of an efficient equilibrium. We extend Krishna (2003)'s framework to accommodate the asymmetric information structure and study how English auction can overcome the informational gap among bidders and achieve the efficient allocation, as opposed to the inefficiency of the sealed bid auction. This extension yields novel insights because we have a natural way of varying insider information by switching an outsider to an insider and establish the linkage principle connecting insider information to the seller's revenues.

Our experimental findings contribute to the experimental literature of auctions that investigates the effects of auction formats on outcomes and bidding behaviour. Most work in the experimental literature have focused on either private value auctions or pure common value auctions. For excellent surveys, see Kagel & Levin (1995) and Kagel & Levin (2011). There are a handful of experimental work on interdependent value auctions. Goeree & Offerman (2002) report an experiment on a first-price auction with interdependent values composed of private and common values and measure the degree of inefficiency by varying the level of competition and the degree of uncertainty on common value information. Kirchkamp & Moldovanu (2004) study experimentally an interdependent value environment with asymmetric bidders in valuation in the English auction and the second-price auction. They found that the English auction yields higher efficiency than the second-price auction, consistent with equilibrium predictions. Boone, Chen, Goeree & Polydoro (2009) study an auction environment with a restricted structure of interdependent values and a single insider in which both English and second-price auctions are inefficient, and report an experimental evidence that the English auction performs better in both efficiency and revenues than the second-price auction. We add novel evidence on the efficiency and revenue performances of the two standard auctions in a general environment with interdependent values and insider information.

The rest of the chapter is organized as follows. In Section 4.2, we develop the model of interdependent value auctions with asymmetric information structure and provide the theoretical results for the second-price and English auctions. Section 4.3 describes the experimental design and procedures. Section 4.4 summarizes experimental findings and Section 4.5 concludes.

4.2 Theory

4.2.1 Setup

Following Kim (2003) a seller has a single, indivisible object to sell to n bidders. Let $N = \{1, \cdots, n\}$ denote the set of bidders. The value of the object to each bidder is determined by n-dimensional information $s \in [0,1]^n$, which we call a signal profile. At this point, we do not specify who observes what signals, which is a central part of our asymmetric information structure and will be discussed shortly. But we adopt the convention of calling i^{th} signal, s_i, bidder i's signal. To denote signal profiles, let $s = (s_j)_{j \in N}$, $s_{-i} = (s_j)_{j \neq i}$, and $s_B = (s_j)_{j \in B}$ for any subset of bidders $B \subset N$. It is assumed that the distribution of signal profile has a full support on $[0,1]^n$. Each bidder i's value, denoted $v_i(s)$, is assumed to be additively separable into two parts, a private value component $h_i : \mathbb{R} \to \mathbb{R}$ and a common value component $g : \mathbb{R}^n \to \mathbb{R}$, i.e.

$$v_i(s) = h_i(s_i) + g(s), \tag{4.1}$$

where $\frac{dh_i}{ds_i} > 0$ and $\frac{\partial g}{\partial s_j} \geq 0$ for all $j \in N$. According to this function, each bidder i's private value component only depends on his own signal s_i while the common value component, which is the same across all bidders, depends on the entire signal profile s. We adopt this functional form in part because it provides us with a natural model of the asymmetric information structure in that some bidders often have superior information about the common value aspect of the object than others.[3] This functional form will also be used later for our experimental study.[4] We assume that h_i and g are twice continuously differentiable and $h_i(0) = g(0) = 0$ for normalization. Assume also that for each i, $\lim_{s_i \to \infty} v_i(s_i, s_{-i}) = \infty$ for any $s_{-i} \in [0,1]^{n-1}$.[5] Note that the value function defined in (4.1) satisfies the *single crossing property*: for all s and $i \neq j$, $\frac{\partial v_i}{\partial s_i} > \frac{\partial v_j}{\partial s_i}$. Note also that for any s, $v_i(s) > v_j(s)$ if and only if $h_i(s_i) > h_j(s_j)$, i.e. a bidder with a higher private component has a higher value. The allocation which gives the object to a bidder with the highest value–or highest private component–for every realization of signal profile in the support is called (ex-post) *efficient*.

The asymmetric information structure is modeled by partitioning N into I, a set of *insiders*, and O, a set of *outsiders*: Each outsider $i \in O$ only knows the private value

[3] There are some other studies which adopt this value function in the standard information setup. For instance, see Wilson (1998) and Hong & Shum (2003)

[4] We note that our theoretical results are not restricted to the assumed functional form. Refer to Kim (2004) for more general condition for the value function under which the efficiency and revenue results of this chapter continue to hold.

[5] This can always be satisfied by, if necessary, redefining the function v_i for signal profiles that are not in the support $[0,1]^n$.

component $h_i(s_i)$ while each insider $i \in I$ knows both the private and common value components, $h_i(s_i)$ and $g(s)$. This assumption is tantamount to having each outsider i be informed of his own signal s_i (due to the monotonicity of h_i). So, if $O = N$ (i.e. there is no insider), then the model boils down to the *standard information case* where every bidder i is only informed of his own signal s_i. Note that we are making a parsimonious assumption on insiders' information by only saying that they know the values of $h_i(s_i)$ and $g(s)$ for each s, while being silent on what signals precisely they are informed of.[6] We impose no restriction on the number of insiders or outsiders except that there are at least one insider and one outsider, i.e. $|I| \geq 1$ and $|O| \geq 1$. The information structure described so far is assumed to be common knowledge among bidders In particular, outsiders know who insiders are. This assumption is reasonable in the examples mentioned in the introduction since it is commonly known: who owns the neighboring tract in an OCS auction; who are the existing shareholders or current management trying to buy the target firm in an takeover auction; who are the expert bidders in an auction of artworks.

For the auction format, we focus on the second-price (sealed-bid) auction and English auction.[7] In the second-price auction, the highest bidder wins the object and pays the second highest bid. For English auction, we consider the Japanese format, where bidders drop out of the auction as the price rises continuously starting from zero until only one bidder remains and is awarded the object at the last drop-out price. (Ties are broken uniform randomly in both the second-price and English auctions.) Despite the similarity in the pricing rules, the two auction formats are different in that in English auction, bidders, especially outsiders, are given chances to observe others' drop-out prices and update their information while they are not in the second-price auction. In both auction formats, we assume that each insider, who knows his value precisely, employs the weakly dominant strategy of bidding (or dropping out at) his value.

4.2.2 Second-Price Auction

Let us start with the analysis of the second-price auction with two bidders, one insider and one outsider. In this case, the analysis of the two auction formats is no different since English auction ends as soon as one bidder drops out so bidders have no chance to update their information as in the second-price auction. The following proposition establishes the efficiency of both auction formats with two bidders.

[6] In the first-price auction, however, what insiders know beyond their values can be important since it provides useful information about their opponents' bids.

[7] We do not consider the first-price auction mainly due to its analytical intractability under the asymmetric information structure. In the case of one insider and one outsider, however, an inefficiency result can be established. Refer to Kim & Che (2004) for this result.

Theorem 4.2.1 (Efficiency with Two Bidders). *With $n = 2$, the object is efficiently allocated in the unique undominated (Bayesian Nash) equilibrium of the second-price (or English) auction.*

The uniqueness of equilibrium follows from the fact that the insider using the undominated strategy pins down the optimal strategy for the outsider, which turns out to be identical to the well-known equilibrium strategy in the standard information case with two bidders (Maskin 1992). The uniqueness and efficiency of the equilibrium stand in contrast with the standard information case where there exist a plethora of inefficient (undominated) equilibria for the second-price auction.[8]

With more than two bidders, however, the second-price auction ceases to be efficient, as the following result shows. The proofs of this theorem and all subsequent results are contained in Appendix 4.6.1.

Theorem 4.2.2 (Inefficiency of Second-Price Auction). *Suppose that $n \geq 3$ and $\frac{\partial g}{\partial s_i} > 0$ for all i. Suppose also that the efficient allocation requires insiders to obtain the object with a positive probability less than one. Then, there exists no efficient equilibrium for the second-price auction.*

To illustrate what causes this inefficiency, consider a symmetric case with 3 bidders where $h_i = h$ for some h and all i, so a bidder which the highest signal has the highest value. Suppose that bidder 3 is an insider while bidder 1 is an outsider. Fix bidder 1's signal at $\bar{s}_1$. In Figure 4.1 below, E_i denotes a set of signal profiles (s_2, s_3) for which bidder i has to be a winner according to the efficient allocation.
Along the line $\overline{CC'}$, bidder 1's bid remains the same. The same has to be true for bidder 3's bid in order that bidder 1 (resp. bidder 3) can be a winner below (resp. above) the line, as the efficiency requires. This is not possible, however, since the value bidding strategy of bidder 3 depends on s_2 as well as s_1 or s_3 so it varies along $\overline{CC'}$.

Intuitively, the above inefficiency is caused by a mismatch between the bidding strategies of outsider and insider in that the former's bid depends only on his own signal while the latter's on the entire signal profile. One might ask why the same intuition does not apply to the two bidder case where the efficiency always obtains (as shown in Theorem 4.2.1). To explain, we need to consider what extra information an insider's bid reflects, compared to the standard case. In the two bidder case, the extra information is a single outsider's signal, which is also known to that outsider. In case of three or more bidders, however, the extra information includes signals unknown to outsiders. For instance, in the above example, bidder 3's bid reflects s_1 and s_2 in

[8] To see this, refer to Bikhchandani & Riley (1991) for the common value case and Chung & Ely (2000) for the interdependent value case.

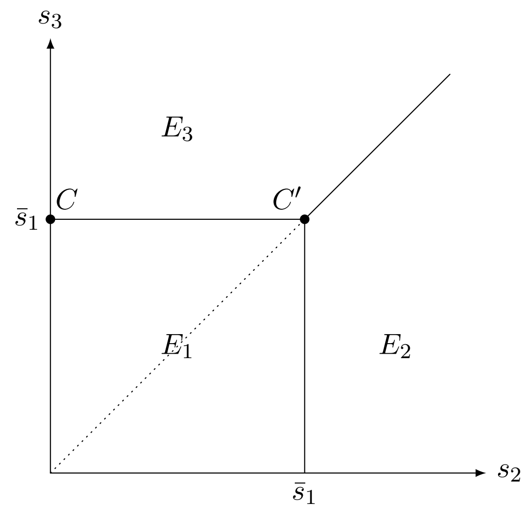


Figure 4.1: Example: Symmetric $I = \{3\}, O = \{1\}$

addition to s_3, and s_2 is not known to outsider bidder 1. The inability of bidder 1 to adjust his bid depending on s_2, in contrast to bidder 3's ability to do so, is what leads to an inefficient allocation in the second-price auction.

The equilibrium bidding strategy for outsiders in the second-price auction is difficult to characterize due to the asymmetry between outsider's and insider's strategies. In the following example, we provide partial characterizations of equilibrium strategies with linear value function that will be used in our experimental study.

Example 1. Suppose that there are three bidders each of whom has a signal uniformly distributed in the interval $[0, 1]$, and that for each $i \in N = \{1, 2, 3\}$, $v_i(s) = as_i + \sum_{j \neq i} s_j$ with $a > 1$. (For the experimental study, we will set $a = 2$.) Let us consider three information structures, $I = \emptyset$, $I = \{3\}$, and $I = \{2, 3\}$. Assuming that insiders use the dominant strategy of bidding their values, we aim to find symmetric Bayesian Nash equilibrium bidding strategy for outsiders, denoted $B : [0, 1] \to \mathbb{R}_+$.

In the case of $I = \emptyset$, Milgrom & Weber (1982a) give us the symmetric equilibrium bidding strategy as follow:

$$B(s_i) = \mathbb{E}_{s_{-i}}[v_i(s_i, s_{-i}) | \max_{j \neq i} s_j = s_i] = (a + \frac{3}{2})s_i.$$

In the case of $I = \{2, 3\}$, the equilibrium bidding strategy for bidder 1, which is a best response to the value bidding by bidder 2 and 3, is given as

$$B(s_1) = \begin{cases} \frac{2a^2+4a+1}{2a+1} s_1 & \text{if } s_1 \in [0, \frac{2a+1}{2a+2}] \\ (2a+3)s_1 - a - 1 & \text{otherwise.} \end{cases} \tag{4.2}$$

The detailed analysis for this result is provided in the Supplementary Material. In Figure 2 below, we reproduce Figure 1 with $\bar{s}_1 = 0.7$ and illustrate how the object is allocated according to the equilibrium bidding strategy given in (4.2). The kinked, dashed line corresponds to the signals (s_2, s_3) for which the equilibrium bid of bidder 1 with $\bar{s}_1 = 0.7$ is equal to the higher of value bids by bidder 2 and bidder 3. So, bidder 1 is a winning (losing) bidder below (above) that line. This implies that in the shaded area A_i, the object is allocated to bidder i, even though his value is not the highest.

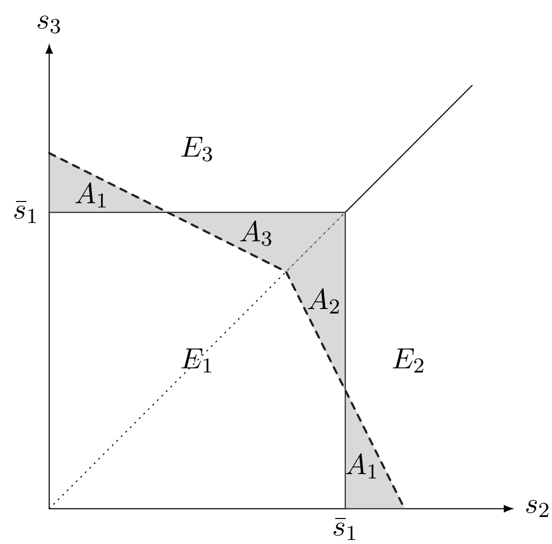


Figure 4.2: Example: Inefficient allocations

In the case of $I = \{3\}$, we only have a partial characterization of monotonic equilibrium bidding strategy that is symmetric for the two outsiders:[9]

Proposition 4.2.3. *For the second-price auction with $I = \{3\}$, any symmetric, strictly increasing equilibrium strategy for outsiders must satisfy the following properties: for $i = 1, 2$, (i) $B(s_i) \in [(a+1)s_i, (a+2)s_i], \forall s_i \in [0, 1]$; and (ii) $B(1) = v_i(1, 1, 1) = a + 2$.*

The Supplementary Material provides the proof of this result along with a numerical

[9] We are *not* claiming here that any symmetric equilibrium must be monotonic. We note that the presence of insider prevents us from establishing the necessity of monotonicity by using the usual argument based on incentive compatibility.

analysis of the equilibrium bidding strategy.

4.2.3 English Auction

In this section, we provide the equilibrium of English auction that achieves the efficient allocation. Using this equilibrium construction, we also establish the comparative statics for the revenue of English auction with respect to the number of insiders. As mentioned earlier, English auction has an advantage over the second-price auction in terms of the additional information bidders can acquire during the dynamic bidding process. We show that this feature enables English auction to overcome the asymmetric information between insiders and outsiders and achieve the efficient allocation. It is also shown that more bidders being insider is beneficial for the seller's revenue.

To this end, we construct the equilibrium of English auction following Milgrom & Weber (1982*a*) and Krishna (2003). A key feature of our equilibrium construction consists of describing how outsiders infer others' signals from their drop-out prices and then how to use this information to determine their own drop-out prices. Assume for the moment that each outsider's drop-out price at any moment of auction is strictly increasing in his signal, so his signal is revealed as he drops out. Also, insiders' drop-out prices are equal to their values. Using this information, the active outsiders who have yet to drop out calculate the *break-even signals* at each current price, which is defined to be a signal profile that makes all active bidders break even if they acquire the object at the current price. Then, each active outsider stays in (exits) the auction if his break-even signal at the current price is smaller (greater) than his true signal. Note that the assumed monotonicity of each outsider's drop-out price with respect to his signal is ensured if his break-even signal as a function of the current price is strictly increasing.

To formalize this idea, let us introduce a few notations. Let A denote the set of active bidders. So, $N\backslash A$, $O\backslash A$, and $I\backslash A$ denote the set of inactive bidders, inactive outsiders, and inactive insiders, respectively. With p_i denoting a drop-out price of bidder i, let $p_B = (p_i)_{i\in B}$ for a subset of bidders $B \in N$. Note that a price profile, $p_{N\backslash A}$, corresponds to a history of the game at a point where only bidders in A are active. Suppose now that the current price is equal to p with a history $p_{N\backslash A}$. Suppose also that the signals of inactive outsiders have been revealed to be $s_{O\backslash A}$. Given these revealed signals, we define the break-even signal profile, denoted $(s_A(p;p_{N\backslash A}), s_{I\backslash A}(p;p_{N\backslash A}))$, to solve the following system of equations:

$$v_i(s_{O\backslash A}, s_A(p;p_{N\backslash A}), s_{I\backslash A}(p;p_{N\backslash A})) = \begin{cases} p_i \ \text{ for } \ i \in I\backslash A & \text{(4.3a)} \\ p \ \ \text{ for } \ i \in A & \text{(4.3b)} \end{cases}$$

The equations in (4.3a) says that with the signal profile $(s_{O\backslash A}, s_A(p; p_{N\backslash A}), s_{I\backslash A}(p; p_{N\backslash A}))$, the value of each inactive insider i is equal to his drop-out price p_i, which is consistent with the insiders' value bidding strategy. The equations in (4.3b) says that if the active bidders acquire the object at the current price p, then they would break even, which is why the solution of (4.3) is called break-even signals. Then, the outsiders' equilibrium strategy is simple: After any history $p_{N\backslash A}$, each outsider i drops out (stays in) at price p if and only if $s_i \leq (>) s_i(p; p_{N\backslash A})$. So, if $s_i(p; p_{N\backslash A})$ is strictly increasing and continuous with p, then an outsider i with signal s_i drops out at price p at which $s_i = s_i(p; p_{N\backslash A})$. It means that each outsider's signal is revealed via his drop-out price, as we assumed in order to set up the system of equations (4.3). The following theorem proves that the value function in (4.1) admits a strictly monotonic solution for the break-even signals, and thereby establishes the existence of efficient equilibrium:

Theorem 4.2.4 (Efficiency of English Auction)**.** *Consider English auction with $n \geq 3$.*

1. *There exists an ex-post Nash equilibrium where each outsider $i \in O\backslash A$ drops out at price p after history $p_{N\backslash A}$ if and only if $s_i < s_i(p; p_{N\backslash A})$, where $s_i(p; p_{N\backslash A})$ solves* (4.3)*.*

2. *The equilibrium bidding strategy described in 1 leads to the efficient allocation.*

An intuitive understanding of the efficiency of English auction can be gained from comparison with the second-price auction. For this, we revisit the three bidder example. Recall from Figure 1 (or Figure 2) that the source of the inefficiency result with the second-price auction is that the insider bidder 3 adjusts his bid according to s_2 (in particular, along the threshold $\overline{CC'}$) while the outsider bidder 1 fails to do so. This problem disappears in English auction. To see it, suppose that bidder 3 is the only insider, and also that bidder 2 drops out first and reveals his true signal, which bidder 1 learns and subsequently reflects in his drop-out strategy. So there no longer is the informational asymmetry between bidder 1 and bidder 3 regarding s_2. Also, from this point on, the bidding competition reduces to the two bidder case with one outsider and one insider as in Theorem 4.2.1, where the efficiency easily obtains. In the current example with three bidders and one insider, the argument for the efficiency result based on the above intuition can be completed by observing that the outsiders' drop-out strategy implies: (i) they drop out before the price reaches their values; (ii) they drop out in order of their values. Due to (i), an insider with the highest value always becomes a winner. In the case an outsider has the highest value, the efficiency is also achieved because, due to (ii), another outsider drops out and reveals his signal before the highest-value outsider does so. The proof of Theorem 4.2.4 generalizes this argument to the case in which there can be more insiders or outsiders. A difficult part of the proof lies in proving the strict monotonicity of the break-even signals, which follows from the fact that the value functions in (4.1) satisfy a sufficient condition provided by Krishna (2003).

Turning to the revenue property, we ask how the presence of more insiders affects the seller's revenue. To do so, we consider two English auctions, E and E', which only differ by one bidder who switches from an outsider in E to an insider in E'. Given a signal profile s, let $P(s)$ and $P'(s)$ denote the seller's revenue in E and E', respectively, under the equilibrium described earlier.

Theorem 4.2.5 (Linkage Principle)**.** *For any signal profile $s \in [0,1]^n$, $P(s) \leq P'(s)$.*

This result follows from establishing two facts: (i') the switched insider drops out at a higher price in E' than in E; (ii') the higher drop-out price of the switched insider causes (active) outsiders to also drop out at higher prices in E'. To provide some intuition behind (ii'), let us revisit the three bidder example. Suppose that bidder 1 is a winner and pays the bidder 2's drop-out price p_2 after bidder 3, the only insider, has first dropped out at p_3. Then, the break-even signals $s_1(p_2;p_3)$ and $s_3(p_2;p_3)$ satisfy

$$v_3(s_1(p_2;p_3), s_2, s_3(p_2;p_3)) = p_2 > p_3 = v_3(s_1, s_2, s_3). \tag{4.4}$$

Since bidder 1 is active at p_2, we have $s_1 > s_1(p_2;p_3)$, which implies by (4.4) that $s_3 < s_3(p_2;p_3)$. The fact that $s_1 > s_1(p_2;p_3)$ means the break-even signal of the active outsider underestimates his true signal, which has the effect of lowering the selling price p_2. However, this effect is mitigated by the fact that $s_3 < s_3(p_2;p_3)$ or the insider's signal is overestimated. The underestimation of active outsiders' signals results from their attempt to avoid the winner's curse by bidding as if the currently unknown signals are just high enough to make them break even at the current price. This is why an outsider drops out before the price reaches his value, as the fact (i') above suggests. Thus, one can say that an outsider getting better informed to become an insider alleviates the detrimental effect of the winner's curse on both his own and other outsiders' drop-out prices.

The revenue result in Theorem 4.2.5 is reminiscent of the linkage principle of Milgrom & Weber (1982*a*) (henceforth MW) in that the shift from E to E' increases the information available to the bidders, There are some crucial differences, however: Unlike MW's, the increased information here is not public since only the switched insider gets better informed. Also, the revenue increase in Theorem 4.2.5 holds for every realization of signal profile and thus does not rely on the assumption that the signals are affiliated.

Example 2. Let us consider the same linear example as in Example 1. In the case of $I = \emptyset$ (i.e., no insider), the equilibrium strategy in Theorem 4.2.4 takes the same form as in MW: if no one has dropped out, then the break-even signal for each bidder i is given as $\frac{p}{(a+1)}$, which means bidder i drops out at price equal to $(a+2)s_i$; if one bidder j has already dropped out at p_j and thereby revealed his signal s_j, then the break-even signal for each remaining bidder i is given as $\frac{1}{a+1}(p - s_j)$, which means bidder i drops out at price equal to $(a+1)s_i + s_j$.

Let us turn to the case $I \neq \emptyset$. If no one has dropped out or one outsider has dropped out, an (active) outsider's drop-out strategy remains the same as above. After an insider $j \in I$ has dropped out at price p_j, the condition in (4.3) becomes

$$\begin{aligned} as_i(p;p_j) + \sum_{k \neq i} s_k(p;p_j) &= p \text{ for each } i \neq j \\ as_j(p;p_j) + \sum_{k \neq j} s_j(p;p_j) &= p_j, \end{aligned}$$

which yields the break-even signal $s_i(p;p_j)$ for each $i \neq j$ that solves

$$(a+1)s_i(p;p_j) + \frac{1}{a}(p_j - 2s_i(p;p_j)) = p.$$

Given the equilibrium strategy that calls for each outsider i to drop out at price p where $s_i(p;p_j) = s_i$, the above equation implies that an outsider i drops out at price equal to $(a+1)s_i + \frac{1}{a}(v_j(s) - 2s_i)$.

Table 4.1 and 4.2 below summarize the drop-out prices that result from the equilibrium strategy described above in the case of $I = \{3\}$ and $I = \{2,3\}$, respectively. Note that we restricted attention to the cases of $s_1 > s_2$ in Table 1 and $s_2 > s_3$ in Table 2, so the value ranking within outsiders/insiders, and thus the order of their drop-out prices, is fixed. The drop-out prices are for the bidders who do not have the highest value, meaning that a bidder with the highest value is winning. The rightmost columns of the two tables show what would be the sale prices (or the second drop-out prices) if there were one less insider, which are lower than the prices in the second column from the right, as one can check easily. Thus, the seller's revenue gets (weakly) higher for each realized signals as an outsider switches to an insider.

4.3 Experimental Design

The experiment was run at the Experimental Laboratory of the Centre for Economic Learning and Social Evolution (ELSE) at University College London (UCL) between

Table 4.1: Drop-out prices for English auction in case of $I = \{3\}$ and $s_1 > s_2$

	1st drop-out price	2nd drop-out price (= the sale price)	the sale price with $I = \emptyset$
(1-i) $s_3 > s_1 > s_2$	$p_2 = (a+2)s_2$	$p_1 = (a+1)s_1 + s_2$	$(a+1)s_1 + s_2$
(1-ii) $s_1 > s_3 > s_2$	$p_2 = (a+2)s_2$	$p_3 = v_3(s)$	$(a+1)s_3 + s_2$
(1-iii) $s_1 > s_2 > s_3$ and $(a+2)s_2 > v_3(s)$	$p_3 = v_3(s)$	$p_2 = (a+1)s_2 + \frac{1}{a}(v_3(s) - 2s_2)$	$(a+1)s_2 + s_3$
(1-iv) $s_1 > s_2 > s_3$ and $(a+2)s_2 < v_3(s)$	$p_2 = (a+2)s_2$	$p_3 = v_3(s)$	$(a+1)s_2 + s_3$

Table 4.2: Drop-out prices for English auction in case of $I = \{2, 3\}$ and $s_2 > s_3$

	1st drop-out price	2nd drop-out price (= the sale price)	the sale price with $I = \{3\}$
(2-i) $s_1 > s_2 > s_3$	$p_3 = v_3(s)$	$p_2 = v_2(s)$	$\max\{v_3(s), (a+1)s_2 + \frac{1}{a}(v_3(s) - 2s_2)\}$
(2-ii) $s_2 > s_1 > s_3$ and $(a+2)s_1 > v_3(s)$	$p_3 = v_3(s)$	$p_1 = (a+1)s_1 + \frac{1}{a}(v_3(s) - 2s_1)$	$(a+1)s_1 + \frac{1}{a}(v_3(s) - 2s_1)$
(2-iii) $s_2 > s_1 > s_3$ and $(a+2)s_1 < v_3(s)$	$p_1 = (a+2)s_1$	$p_3 = v_3(s)$	$v_3(s)$
(2-iv) $s_2 > s_3 > s_1$	$p_1 = (a+2)s_1$	$p_3 = v_3(s)$	$v_3(s)$

December 2011 and March 2012. The subjects in this experiment were recruited from an ELSE pool of UCL undergraduate students across all disciplines. Each subject participated in only one of the experimental sessions. After subjects read the instructions, the instructions were read aloud by an experimental administrator. Each experimental session lasted around two hours. The experiment was computerized and conducted using the experimental software z-Tree developed by Fischbacher (2007). Sample instructions are reported in Appendix 4.6.4.

In the design, we use a variety of auction games with three bidders, $i = 1, 2, 3$. Each bidder i receives a private signal, s_i, which is randomly drawn from the uniform distribution over the set of integer numbers, $\{0, 1, 2, ..., 100\}$. Given a realization of signal profile $s = (s_1, s_2, s_3)$, the valuation of the object for each bidder i is

$$v_i(s) = 2 \times s_i + \sum_{j \neq i} s_j.$$

We have two auction formats in the experiment—the second price (sealed-bid) auction and the English auction. In each auction format, we had three distinct games in terms of the number of insiders from zero to two, $k = 0, 1, 2$. Thus, there are in total six treatments in terms of the auction format and the number of insiders. A single treatment consisting of one auction format and a single value of k was used for each session. We conducted 12 sessions in total by having two sessions for each treatment of the auction games. Each session consisted of 17 independent rounds of auction games, while the first two rounds were practice rounds in which auction outcomes were not counted for actual payoffs. Throughout the chapter, we use the data only after the first two rounds. The table 4.3 below summarizes the experimental design and the amount of experimental data. The first number in each cell is the number of subjects and the second is the number of group observations in each treatment. In total, 233 subjects participated in the experiment.

Table 4.3: Experimental sessions

Auction format	# of insiders (k)	Session 1	Session 2	Total
English	0	21 / 105	21 / 105	42 / 210
	1	18 / 135	20 / 150	38 / 285
	2	23 / 345	19 / 285	42 / 630
Second-price	0	21 / 105	21 / 105	42 / 210
	1	20 / 150	16 / 120	36 / 270
	2	17 / 255	16 / 240	33 / 495

We use an irrevocable-exit, ascending clock version of the English and the second-price auctions (Kagel, Harstad & Levin 1987, Kirchkamp & Moldovanu 2004). In the English auction treatments, each subject sees three digital clocks representing the bidding process on his or her computer screen, one for each bidder in the group, while there is only one clock presented for each subject on the screen in the second-price auction treatments. In the English auction, the computer screen clearly indicates which clock belongs to each subject and, if any, belongs to an insider.

In the beginning of an auction round, the subjects were randomly formed into three-bidder groups. In treatments with insider(s), each insider bidder was played by the computer, while outside bidders were played by human subjects. The groups formed in each round depend solely upon chance and were independent of the groups formed in any of the other rounds. Once assigned into a three-bidder group, each subject observed the realization of his or her own priviate signal and the formula of the valuation computation. Other bidders' signals in the formula were hidden to indicate that subjects could not observe others' signals.

Clocks simultaneously start at -4 and synchronously move upwards by 1 unit per half second. In the second-price auction, when the subject drops out by stopping his or her clock, it turns red with the price at which it was stopped. If the other two participants have not dropped out yet, their clocks continue to ascend by the same speed. Once all three bidders have stopped their own clocks, the auction is then over. The last bidder who chose the highest price becomes a winner of the auction and pays the price at which the second bidder dropped out. On the other hand, in the English auction, if one bidder stops his or her clock, the remaining two bidders observe, at the next bid increment, that bidder's clock having been stopped and turning red. There will then be 3 seconds of time pause before the two remaining clocks will synchronously increase by 1 unit per second. If one more participant drops out, the auction will then be over. The last remaining bidder becomes a winner of the auction and pays the price at which the second bidder dropped out. If all remaining bidders dropped out at the same price level or if the price level reached 500 (the maximum bid allowed), the winner is then selected at random from the set of those participants and pay the price at which this event occurred.

Each subject simply needs to move the mouse over his or her own clock and click on it, when the price on the clock reaches the level he or she wants to drop out. This makes the subject drop out of the bidding, that is, his or her own clock stop. Once subjects have dropped out, they were not allowed to re-enter the auction. Subjects cannot stop their clocks before it reaches 0, which is the minimum bid allowed.

In treatments with at least one insider, an insider is played by the computer. This (computer-generated) insider uses a simple rule of drop-out decision: it drops out at a price equal to its own valuation. The computer participant always abides by this rule. This information is common knowledge to subjects.

When an auction round ended, the computer informed each subject of the results of that round, which include bids at which bidders dropped out, signals that bidders received, values of the auction object, payments and earnings in that round. Once every subject confirmed the results, the next round started with the computer randomly forming new groups of three bidders and selecting signals for bidders. This process was repeated until all 17 rounds were completed.

Earnings were calculated in terms of tokens and then exchanged into pounds at the rate where each 40 tokens was worth £1. The earnings in each round were determined by the difference between winning revenue and winning cost. The winning revenue is the valuation assigned to the subject if he or she won the auction and zero otherwise. The winning cost is the price paid by the subject. If the subject did not win the auction, the winning cost is simply equal to zero. In our experiment, subjects may accumulate losses, exhaust their balance, and go bankrupt during the experiment. If this event happened, subjects were no longer allowed to participate in the experiment. In order to avoid the potentially adverse impacts of the limited liability for losses on bidding behaviour,[10] we gave each subject a relatively large amount of money, £10, as an initial balance that the subjects uses in the experiment. This initial balance is only paid to subjects in the case they actually earn it in the auction games. For instance, if their net earning in the auction games is zero, then they only receive the participation fee, 5 pounds. None of the subjects experienced a bankruptcy during the experiment. The total payment to a subject was the sum of his or her earnings over the 15 rounds after the first two practice rounds, plus the initial balance £10 and an extra £5 participation fee. The average payment was about £18.32. Subjects received their payments privately at the end of the session.

4.4 Experimental Findings

4.4.1 Efficiency

We begin the analysis of experimental data by comparing the efficiency performance across auction treatments. Table 4.4 reports the frequency of efficient allocation as well as the average efficiency ratio measuring the economic magnitude of inefficient outcomes. The efficiency ratio is defined as the actual surplus improvement over random allocation as percentage of the first-best surplus improvement over random assignment.[11] This ratio equals one if the allocation is efficient, that is, the highest-value bidder wins the object, and less than one otherwise. For auction treatments with one insider (resp. two insiders), we divide the data with respect to the ranking of the value of the insider (resp. the outsider). The last two columns on the right side report, for each treatment of the number of insiders, p-values from the one-sided t-test of the null hypothesis that the efficiency outcome of the English auction is less than or equal to that of the second-price auction.

The level of efficiency is high in all treatments. The frequencies of efficient allocation range from 76% (in the second-price auction with $k = 1$) to almost 90% (in the English

[10]Hansen & Lott (1991) argued that the aggressive bidding behaviour in a common value auction experiment done by Kagel & Levin (1986) may be a rational response to the limited liability rather than a result of the winner's curse. Lind & Plott (1991) designed an experiment eliminating the limited-liability problem and found that this problem does not account for the aggressive bidding patterns in the Kagel and Levin experiment.

[11]Our measure of efficiency ratio normalizes the realized surplus both by the best-case scenario (efficiency) and by the worst-case scenario (random assignment). This double-normalization renders the measure more robust against the rescaling of value support than an alternative measure like the percentage of the first-best surplus realized.

Table 4.4: Frequencies and ratios of efficient allocation

	SPSB		English			
	(1)	(2)	(3)	(4)	$H_0 : (1) = (3)$	$H_0 : (2) = (4)$
# of insiders	Freq.	Ratio	Freq.	Ratio		
0	0.78	0.80	0.83	0.84	0.17	0.22
1	0.76	0.74	0.83	0.89	0.05	0.00
2	0.83	0.86	0.89	0.93	0.01	0.00

Notes: The efficiency ratio is defined as (realized surplus minus random surplus) divided by (first-best surplus minus random surplus). The two columns on the right side report the p-value from the t-test for the null hypothesis that outcomes between the SPSB and the English auctions are equivalent.

auction with $k = 2$). The efficiency ratio shows similar patterns: the subjects achieved, on average, between 74% (in the second-price auction with $k = 1$) and 93% (in the English auction with $k = 2$) of the first-best surplus over random surplus. On the other hand, we find that the English auction exhibits better performance in efficiency than the second-price auction, when theory predicts so. In the symmetric information structure with no insider ($k = 0$), there is no significant difference in efficiency between the two auction formats. In the presence of insider, however, the efficiency outcomes are significantly higher in the English auction than in the second-price auction. These results are qualitatively consistent with the theoretical predictions that when at least one insider is present in auction, an efficient equilibrium exists in the English auction but not in the second-price auction. We further divide the data with respect to the value ranking of bidders. In the auctions with one insider ($k = 1$), the efficiency ratio of the English auction is significantly higher than that of the second-price auction, regarding of the value ranking. In the auctions with two insiders ($k = 2$), the English auction attains higher efficiency than the second-price auction in each case of value ranking except for the case where the outsider has the second-highest value.

To examine further the higher efficiency performance of the English auction, we divide the data of the English auction treatments into two subsamples with respect to whether the second-price auction format, if it would have been used, had predicted an efficient or inefficient allocation. That is, for each observed realization of signals in each data point of the English auction treatments, we check if the second-price auction format predicts an inefficient or efficient allocation. Due to the small sample size of the second-price predicting an inefficient allocation, we allow a 5-token margin for classifying the case of inefficiency in such a way that treats a single data point as *inefficient* if the equilibrium of the second-price auction is either inefficient or efficient but the difference of its two high bids is less than 5 tokens. We then check how often subjects were able to achieve an efficient allocation in each of these two cases. We conduct the same analysis for the second-price auction treatments. Table 3 presents the frequencies of efficient allocation in each case in the treatments with at least one insider with the number of observations in parentheses.[12] The last column reports p-values from the one-sided t-test of the null hypothesis that the efficiency outcome of the English auction is less than or equal to that of the second-price auction.

Table 4.5: The decomposition of efficient and inefficient allocations

	Highest-value bidder							
	English auction				SPSB auction			
	Outsider		Insider		Outsider		Insider	
Winner	Efficient	Inefficient	Efficient	Inefficient	Efficient	Inefficient	Efficient	Inefficient
	4.5-A. (# of insiders) = 1							
Outsider	0.83	0.11	–	0.19	0.78	0.06	–	0.30
	(156)	(21)		(19)	(147)	(11)		(25)
Insider	–	0.05	0.81	–	–	0.16	0.70	–
		(10)	(79)			(30)	(57)	
Total	0.83	0.17	0.81	0.19	0.78	0.22	0.70	0.30
	(156)	(31)	(79)	(19)	(147)	(41)	(57)	(25)
	4.5-B. (# of insiders) = 2							
Outsider	0.87	–	–	0.11	0.72	–	–	0.12
	(180)			(46)	(118)			(39)
Insider	–	0.13	0.89	–	–	0.28	0.88	–
		(26)	(378)			(46)	(292)	
Total	0.87	0.13	0.89	0.11	0.72	0.28	0.88	0.12
	(180)	(26)	(378)	(46)	(118)	(46)	(292)	(39)

Table 4.5 offers further insights on the efficiency performances across treatments with insiders. In the case where the second-price auction format predicts an efficient allocation, actual frequencies of efficiency are quite high, ranging from 81% (in the second-price auction treatment with $k = 1$) to 92% (in the English auction treatment with $k = 2$). In contrast, the level of efficiency becomes much lower in the case where the equilibrium under the second-price auction format is treated as inefficient: the frequencies of efficiency range between 33% (in the second-price auction with $k = 1$) and 66% (in the English auction treatment with $k = 2$). In each given treatment, the difference of efficiency outcomes between these two cases are statistically significant at

[12]The results of Table 4.5 remain basically the same either when we do not use any margin or larger margin in classifying the case where the second-price auction format predicts an inefficient outcome.

usual significance levels. Thus, where the bidding mismatch problem between outsiders and insiders is prone to occur, subjects tend to more likely fail to attain an efficient allocation. Furthermore, in the case where the second-price auction format predicts an inefficient allocation, the English action performs significantly better than the second-price auction when $k = 1$: 33% in the second-price auction *versus* 62% in the English auction. In that case of the treatments with $k = 2$, we find no difference between the two auction formats.

We summarize the efficiency outcomes as follows.

Finding 1 (efficiency) *The English auction exhibits higher efficiency performance than the second-price auction in the presence of insiders, as theory predicts. In the symmetric information structure where there is no insider, there is no of efficiency performance between the two auction formats.*

4.4.2 Revenue

We move on to the comparison of revenue performances across auctions. Table 4.6 presents average percentage deviations of observed revenues from their theoretical predictions across treatments, along with their standard errors and p-value from t-tests for the null hypothesis that the sample mean is equal to zero. For auction treatments with one insider (resp. two insiders), we divide the data with respect to the ranking of the value of the insider (resp. the outsider).

When all bidders are outsiders, we observe that observed revenues are significantly higher than theoretically predicted revenues: 22% higher in the English auction and 26% higher in the SPSB auction. This tendency becomes weaker in auctions with at least one insider: 3% (0%) higher in the English auction with $k = 1$ ($k = 2$) and 14% (1%) higher in the SPSB auction with $k = 1$ ($k = 2$). It may not be unexpected because insiders in the experiment are computer-generated and play the equilibrium strategy of bidding their own values. Despite this consideration, observed revenues in the SPSB auction with $k = 1$ are significantly above theoretical prediction. When we look closer at the data by the ranking of values, in auctions with $k = 1$, the tendency of revenues being above theoretical prediction becomes strong when both outsiders have lower or higher values than the insider. This apparently results from the overbidding (relative to the BNE equilibrium) by outsiders. We will investigate the bidding behaviour of subjects more thoroughly in the next subsection. In auctions with $k = 2$, observed revenues appear to concentrate around theoretical prediction. The magnitude of the departures of observed revenues from theoretical ones is not large, although some of the departures remain significant.

Table 4.6: Average percentage deviations of observed revenues from theoretically predicted revenues

# of insiders	Ranking of values (I or O)	English auction	SPSB auction
0	All	0.22	0.26
		(0.096, 0.023)	(0.045, 0.000)
1	All	0.03	0.14
		(0.014, 0.047)	(0.024, 0.000)
	I = (highest-value)	0.06	0.28
		(0.036, 0.101)	(0.058, 0.000)
	I = (second highest-value)	-0.02	0.03
		(0.012, 0.112)	(0.019, 0.165)
2	I = (lowest-value)	0.03	0.13
		(0.012, 0.005)	(0.036, 0.001)
	All	0.00	0.01
		(0.002, 0.453)	(0.004, 0.063)
	O = (highest-value)	-0.01	-0.03
		(0.003, 0.000)	(0.005, 0.000)
	O = (second highest-value)	0.02	0.03
		(0.006, 0.008)	(0.010, 0.001)
	O = (lowest-value)	0.00	0.02
		(0.001, 0.024)	(0.006, 0.005)

Notes: I stands for an insider and O represents an outsider. The first number in parentheses is a standard error of sample mean and the second number is p-value from t-test for the null hypothesis that the mean is equal to zero.

Our theory establishes the linkage principle of the English auction that for any signal profile, the switch of an outsider to an insider weakly increases the revenue. In order to check out the linkage principle, we run regressions of observed revenues (resp. theoretical revenues) on signal profiles and dummies for the number of insiders in each auction format. The results are reported in Table 4.7.[13]

Controlling for the signal profile, switching an additional outsider to an insider improves revenues significantly in the English auction. The observed revenues increase on average by 10 from $k = 0$ to $k = 1$ and 8 from $k = 1$ to $k = 2$ in the English auction, both of which are statistically significant at usual significant levels. The magnitudes of revenue improving with extra insider in the data are also consistent with those predicted by theory. In the regressions with theoretical revenues, we similarly observe

[13]As a robustness check of Table 4.7, we conduct regression analysis with more flexible functional specifications of quadratic forms of signals or dummies for insider / outsider and their interactions with signals. These results are reported in the Appendix tables 4.11-4.12. In essence, the empirical findings about the linkage principle remain unchanged.

Table 4.7: Regression analysis of revenues

Variables	**English**	**SPSB**
k = 1	10.466***	10.491***
	(1.983)	(2.477)
k = 2	18.716***	9.838***
	(1.738)	(2.218)
s(1)	0.889***	0.766***
	(.039)	(.055)
s(2)	2.087***	1.837***
	(.038)	(.051)
s(3)	0.796***	0.761***
	(.04)	(.054)
constant	-7.739***	18.436***
	(2.98)	(3.883)
# of obs.	1125	975
R2	0.907	0.837
	p-value	
$H_0 : (k = 0) = (k = 1)$	0.000	0.000
$H_0 : (k = 1) = (k = 2)$	0.000	0.749

Notes: Standard errors are reported in parentheses. *, **, and *** represent 10%,5%, and 1% significance level, respectively. $s_{(1)} = min[s]$, $s_{(3)} = max[s]$, $s_{(2)} = med[s]$

the increase in revenues by 12 from $k = 0$ to $k = 1$ and 9 from $k = 1$ to $k = 2$ in the English auction. Analogously, we observe the revenue-improving outcomes in the SPSB auction data, despite that we are unable to prove it theoretically. Observed revenues in the SPSB auction increase by 10 from $k = 0$ to $k = 1$, while they remain unchanged from $k = 1$ to $k = 2$. At least in the current experimental setup, theory predicts the linkage principle in that theoretical revenues increase by 14 from $k = 0$ to $k = 1$ and by 8 from $k = 1$ to $k = 2$.

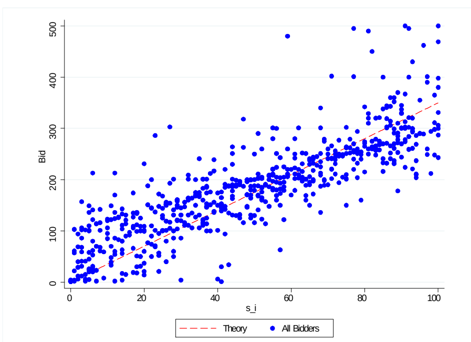
We summarize our findings about revenue as follows.

Finding 2 (revenue) *Revenues in the data tend to deviate above from theoretical prediction, in particular in auctions with no insider. Despite this tendency, the increase in the number of insiders has a positive impact on revenues in the English auction, consistent withe the linkage principle of the English auction. We find similar patterns of revenue-improving with extra insiders in the SPSB auction.*

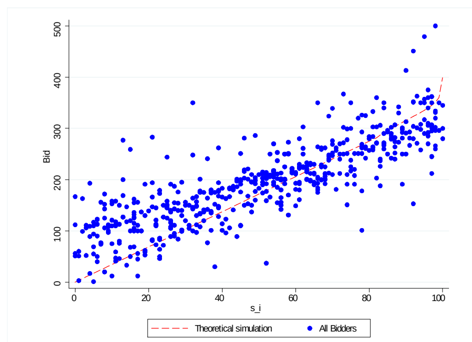
4.4.3 Bidding behaviour

We have found that the auction treatments, in terms of both auction formats and the number of insiders, have significant impacts on efficiency and revenues. At the same time, we have identified some quantitive departures from the predictions of Bayesian Nash equilibrium. In this section, we examine the subjects' behaviour of bidding closely to better understand the features of the data. We begin this analysis with the SPSB auction.

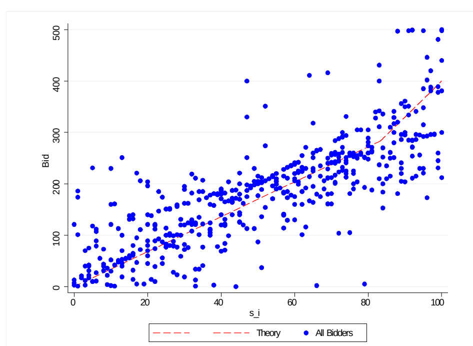
SPSB auction We first overview the general patterns of bidding behaviour in the SPSB auction by drawing scatter plots between subjects' bids and their private signals across insider treatments, and matching them with the BNE strategy. The scatter plots are only based on human participants in the experiment who are outsiders. This is presented in Figure 4.3.



(a) (# of insiders)=0



(b) (# of insiders)=1



(c) (# of insiders)=2

Figure 4.3: Scatter plots of bids and signals: SPSB auction

This simple graphical representation of bidding behaviour already reveals some useful information on the nature of bidding behaviour. There is notable departure of bidding from the BNE strategy when the value of private signal is low: low-signal bidders bid substantially higher than the BNE strategy dictates. This overbidding pattern in low values of signal appears present in all insider treatments, although this pattern appears to be weaker in the treatment with two insiders. On the other hand, subjects' bids appear less responsive to their own private signals than the BNE predicts. As a consequence, many observed bids tend to lie below the BNE strategy when the value of private signal is high.

In order to examine the bidding behaviour of subjects more closely, we run the linear regressions of subjects' bids on their private signals. Our theory predicts that the BNE strategy has a kink at the value of signal equal to 82.109 when $k = 1$ and to 500/6 when $k = 2$. We thus use the regression specifications with and without these kinks.[14] Table 4.8 reports the results of the regressions with robust standard erros clusted by individual subjects. We also present p-values of the F-test for the null hypothesis that observed bids follow the equilibrium strategy.

Table 4.8: Regressions of bids on signals in the SPSB auctions

Variables	**(k = 0)**	**(k = 1)**		**(k = 2)**	
	(1)	(2)	(3)	(4)	(5)
s_i	2.754***	2.423***	2.309***	2.912***	2.756***
	(0.16)	(0.14)	(0.13)	(0.22)	(0.22)
$\mathbb{1}[s_i > 82.109]$			-90.933		
			(79.09)		
$\mathbb{1}[s_i > 82.109]s_i$			1.125		
			(0.89)		
$\mathbb{1}[s_i > 500/6]$					-352.69
					(226.80)
$\mathbb{1}[s_i > 500/6]s_i$					4.006
					(2.50)
Constant	48.797***	72.214***	75.876***	34.475**	39.836***
	(8.57)	(8.12)	(7.29)	(12.78)	(12.07)
R2	0.71	0.7	0.7	0.64	0.65
# of obs.	630	540	540	495	495
$H_0 : (\beta, c) = (b, 0)$	(s=3.5)	n/a	n/a	n/a	$s = \frac{17}{5}, 1[.] = -300, s + 1[.]s = 7$
F test	16.21				3.64
p value	0.00				0.02

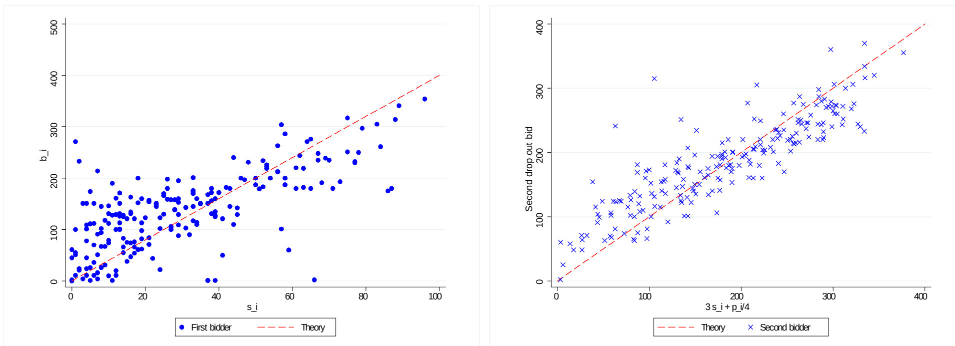
Notes: Robust standard errors clustered by individual subjects are reported in parentheses. *, **, and *** represent 10%, 5%, and 1% significance level, respectively.

[14] As shown in Proposition 1, we only have a partial characterization of the BNE strategy when $k = 1$. Despite that, we are able to numerically derive the kink of the BNE strategy when $k = 1$.

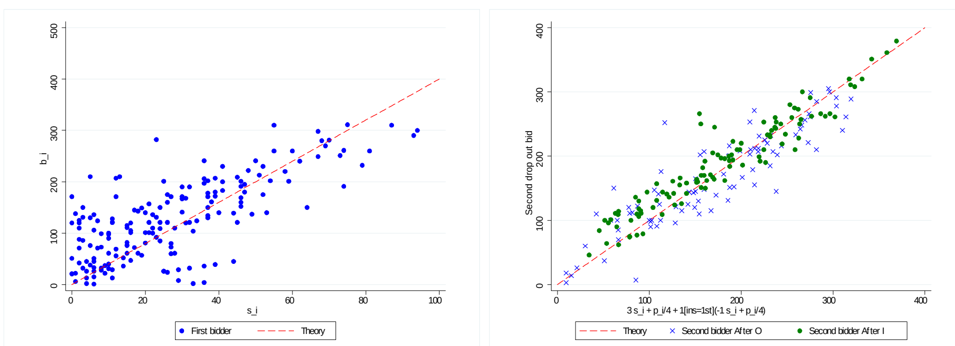
The regression analysis confirms the information extracted from the scatter plots in Figure 4.3. The subjects respond less sensitively to their own private signals than theory predicts. The estimated coefficients on signals are in the range between 2.4 and 2.9, and significantly lower than the theoretical prediction around 3.5 in each insider treatment. The constant term in the regression is significantly positive in each treatment. In sum, the regression results confirm that subjects tend to overbid relative to the equilibrium when signals are low, and that this overbidding tendency reduces when signals are high. The joint test based on F-statistic indicates that subjects' behaviour differ significantly from the equilibrium strategy at usual significance levels. Overall, the overbidding pattern in our data is consistent with the findings in the experimental literature of auctions (Kagel & Levin 2011).

English auction We now turn to the subjects' behaviour in the English auction. We again overview the general patterns of bidding in this auction format by drawing scatter plots with subjects' bids. If the subject is the first drop-out bidder, we relate his drop-out price to his own private signal. If the subject is the second drop-out bidder, we associate his drop-out price with the BNE strategy we constructed in Theorem 3 and showed in Table 1. The second drop-out bidder's equilibrium strategy contains information on her private signal as well as the first drop-out price (and the identity of the first drop-out bid when $k \geq 1$). Thus, it is a convenient, visual way of summarizing the behaviour of second drop-out prices and comparing them with the BNE strategy. The scatter plots are presented in Figure 4.4.

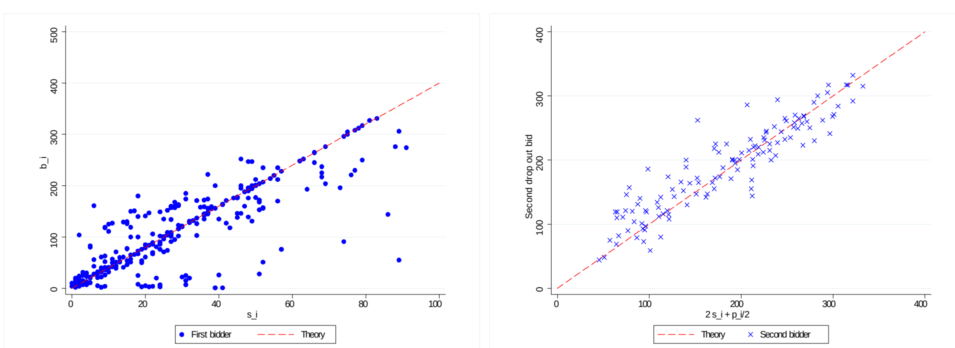
Analogous to the overbidding pattern in the SPSB auction, we observe that the first drop-out prices tend to depart significantly from the equilibrium strategy when the value of private signal is low. However, this pattern appears strikingly less notable in the treatment with two insiders, wherein there is a cluster of observed bids along the line of the equilibrium strategy even when the signal value is quite low. It also appears that the first drop-out subjects tend to respond less sensitively to their private signals. Regarding the second drop-out subjects, they appear to bid less responsively to the combination of their own signal and the first drop-out price than the Bayesian equilibrium dictates.



(a) (# of insiders)=0



(b) (# of insiders)=1

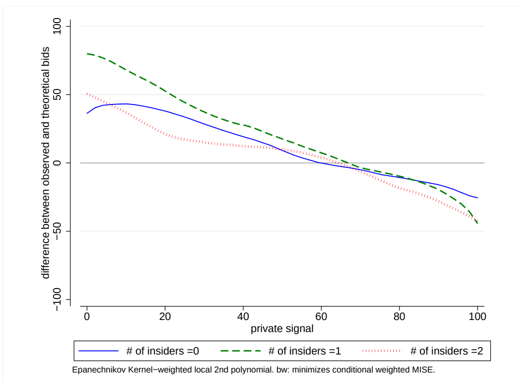


(c) (# of insiders)=2

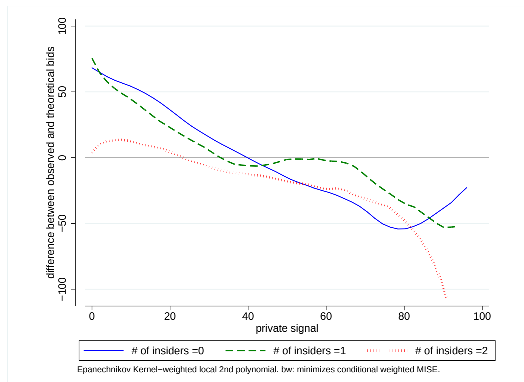
Figure 4.4: Scatter plots of bids and signals: English auction

A useful comparisson between SPSB and English auction bidding behaviour is found in Figure 4.5. We show the Kernel smoothing of the difference between observed and theoretical bids for each treatment. If there were no departures from theory the plotted lines would be horizontal at the 0 value. It makes evident the overbidding (underbidding) at low (high) signals. Increasing the number of insiders has a larger impact on English format, reducing overbidding at low signals.

Figure 4.5: Kernel Estimation of the difference between observed bids and theoretical bids



(a) SPSB auction (all samples)



(b) English auction (first-drop bids)

It is not clear from the simple scatter plot or the non–parametric smoothing whether this behavioural departure results from the insensitivity to their private signal or to the first drop-out price or some other combination.

Therefore, we run the censored regressions of first drop-out and second drop-out prices, with the sample of outsiders. We need to adopt the censored regression method because we only observe the first drop-out price for the lowest bidder and the other two remaining bids are right-censored, and because the second drop-out price is left-censored by the first drop-out price. In the treatment with one insider ($k = 1$), as the BNE strategy predicts, the regression specification for the second drop-out price interacts the second drop-out bidder's private signal and the first-drop price with the dummy indicating if the first drop-out bidder is an insider. We further include an alternative specification by adding this dummy in the regression equation to capture any potential empirical impact of this dummy on the constant term. Table 4.9 reports the regression results and p-values from the F-test for the joint null hypothesis that observed bids follow the equilibrium strategy.[15] Robust standard errors clustered by individual subjects are reported in parentheses.

[15] The censored regression approach, using the maximum likelihood estimation method, is given in detail in Appendix 4.6.3.

Table 4.9: Censored regressions of bids in the English auctions

	(k = 0)		(k = 1)				(k = 2)	
	first drop-out	second drop-out	first drop-out	second drop-out			first drop-out	second drop-out
Variables	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
s_i	2.878***	1.893***	3.258***	2.072***	2.394***	2.456***	3.48***	1.268***
	(0.14)	(0.10)	(0.16)	(0.08)	(0.15)	(0.16)	(0.14)	(0.10)
$\frac{p}{4}$		1.591***		1.639***	1.369***	1.409***		2.812***
		(0.12)		(0.13)	(0.21)	(0.21)		(0.19)
$\mathbb{1}[\text{1st drop-out} \in I]s_i$					-1.010***	-1.076***		
					(0.21)	(0.22)		
$\mathbb{1}[\text{1st drop-out} \in I]\frac{p}{4}$					1.240***	1.093***		
					(0.33)	(0.35)		
$\mathbb{1}[\text{1st drop-out} \in I]$						10.775		
						(8.45)		
Constant	77.893***	46.28***	65.669***	30.733***	22.705***	17.285***	32.089***	21.836***
	(9.61)	(5.60)	(10.13)	(3.63)	(3.82)	(6.28)	(8.04)	(6.79)
σ	61.947	30.677	58.226	29.078	26.358	26.378	56.556	23.396
	(8.365)***	(2.896)***	(6.230)***	(1.760)***	(1.794)***	(1.810)***	(8.449)***	(1.785)***
pseudo-R2	0.13	0.20	0.16	0.22	0.23	0.23	0.16	0.24
# of obs.	630	420	570	388	388	388	630	343
$H_0 : (\beta, c) = (b, 0)$	(s=4)	(s=3,p/4=1)	(s=4)	n/a	$s = 3, \frac{p}{4} = 1,$ $\mathbb{1}_{[.]}s + s = 2,$ $\mathbb{1}_{[.]}\frac{p}{4} + \frac{p}{4} = 2$	$s = 3, p/4 = 1,$ $\mathbb{1}_{[.]}s + s = 2, \mathbb{1}_{[.]}\frac{p}{4}$ $+\frac{p}{4} = 2, \mathbb{1}_{[.]} = 0$	$s = 4$	$s = 2, \frac{p}{4} = 2$
F test	40.17	45.63	21.11		36.65	30.63	9.43	48.41
p value	0.00	0.00	0.00		0.00	0.00	0.00	0.00

Notes: Robust standard errors clustered by individual subjects are reported in parentheses. *, **, and *** represent 10%, 5%, and 1% significance level, respectively.

The regression analysis of the first drop-out prices reveals that the subjects respond to less sensitively to their private signal than the equilibrium strategy across insider treatments. The estimated coefficients on private signal are 2.88 in the treatment with no insider ($k = 0$), 3.26 in the treatment with one insider ($k = 1$), and 3.48 in the treatment with two insiders ($k = 2$), whereas the equilibrium behaviour responds to private signal by a factor of 4. We also found that the constant term is significantly positive in all insider treatments, 77.89 when $k = 0$, 65.67 when $k = 1$, and 32.09 when $k = 2$. Given these results, the null hypothesis that the first drop-out prices follow the BNE strategy is rejected at usual significance levels in each insider treatment. Intriguingly, the estimated coefficient on private signal increases and the constant term declines as the number of insiders increases. Thus, the overbidding pattern and its resulting winner's curse get weaker when there are more insiders in the experiment. The presence of an insider who knows the value of the object may make outsiders be more wary of hedging against the informational asymmetry between insider and outsider. This need to hedge against the informational asymmetry may operate in the direction of correcting the winner's curse.

We turn to the regression analysis of the second drop-out prices, which is quite revealing. Similar to the first drop-out bidders, the second drop-out bidders respond less sensitively to their own private signal than the equilibrium analysis predicts. However, they respond excessively to the first drop-out prices. According to our theory, the equilibrium bid would respond to the first drop-out price by a factor of 0.25 in the treatment with no insider and in the treatment with one insider when an outsider dropped out first, and by a factor of 0.5 in the treatment with one insider when the first drop-out bidder is an insider as well as in the treatment with two insiders. In the experiment, the subjects on average responded to the first drop-out price by about 0.40 when $k = 0$, 0.34 when $k = 1$ and the first drop-out bidder is an outsider, 0.65 when $k = 1$ and the insider dropped out first, and 0.70 in the treatment with two insiders.[16] Despite the excessive responsiveness to the first drop-out price, as suggested in Figure 3, the combined behavioural response of own signal and first drop-out price appears less sensitive that the equilibrium dictates. The joint null hypothesis that experimental behaviour is equivalent to the equilibrium strategy is rejected at usual significance levels in each treatment. Finally, we note that in the auction treatment with one insider ($k = 1$), the subjects responded differentially to the identity of the first drop-out bidder: they tend to put more weight on the first drop-out price and less on their private signal when the insider dropped out first. This is qualitatively consistent

[16]The p-values of the t-test for the null hypothesis that the coefficient of $p/4$ is equivalent to the equilibrium prediction are 0.000 in each case of consideration.

with the Bayesian equilibrium analysis.

Quantifying naive bidding The overbidding pattern in our interdependent value environment with insider information is closely related to the findings of the winner's curse in the experimental literature of common value auctions (see Kagel and Levin, 2002). Winning against other bidders implies that the outsider's value estimate happens to be the highest among outsiders *as well as* higher than each insider's value. Thus, the failure to account for this adverse selection problem results in overbidding and can make the outsider fall prey to the winner's curse.

We employ a simple strategy of quantifying the extent to which subjects in our experiment fail to account properly for the adverse selection problem and thus bid naively. We define naive bidding as bidding based on the unconditional expected value (by ignoring completely the adverse selection problem). For the second-price auction and the first drop-out bidder in the English auction, the naive bidding strategy takes the form of $b^{naive}(s_i) = 2s_i + 100$. We also assume the same form of naive bidding strategy after observing a first-drop price in the English auction, which means that the naive bidder ignores any information from first drop-out price. We then consider a convex combination of the naive bidding strategy and the Bayesian Nash equilibrium strategy. For all bidders in the second-price auction and the first drop-out bidders in the English auction, this combined bidding strategy is represented by

$$b(s_i; \alpha) = \alpha \times b^{naive}(s_i) + (1 - \alpha)\, b^{BNE}(s_i).$$

For the second drop-out bidders in the English auction with first drop-out price p_j, it can be written as

$$b(s_i, p_j; \alpha) = \alpha \times b^{naive}(s_i) + (1 - \alpha)\, b^{BNE}(s_i, p_j).$$

We estimate α by matching this form to the data for each treatment. α measures the degree to which the subjects' behaviour departs from the Bayesian Nash equilibrium and is close to the naive bidding strategy. In our setup, α is well identified. For instance, for the first drop-out bidder in the English auction with each k, this convex combination can be rewritten as $b(s_i; \alpha) = 100 \times \alpha + (2\alpha + 4(1 - \alpha))\, s_i$ and α is identified by matching the constant term and the slope of this equation to the data.

Comparing α estimates across situations where the Bayesian Nash equilibrium is common is of particular interest. Because the equilibrium strategies are held constant in such situations, comparing the magnitude of α estimates is meaningful. For this purpose, we note that the Bayesian Nash equilibrium strategies for the first drop-out bidder are the same in all insider treatments of the English auction. The equilibrium strategies for the second drop-out bidder are common in the English auction with no insider ($k = 0$) and in the English auction with one insider ($k = 1$) when the first drop-out bidder is an outsider. Also, they are common in the English auction with one insider ($k = 1$) when the first drop-out bidder is an insider and in the English auction with two insiders ($k = 2$). We will focus on the comparisons of α in such situations.

Table 4.10 reports the regression results of α estimates across auction treatments. Robust standard error clustered by individual subjects are reported in parentheses. We also report t-test for the equivalence of α between two insider treatments in a given auction format.

Table 4.10: Nonlinear least squares imposing BNE parameters

	SPSB		**English**		
	k=0	k=2	k=0	k=1	k=2
	0.488	0.402	0.657	0.525	0.179
	(.086)***	(.122)***	(.080)***	(.096)***	(0.060)***
R^2	0.94	0.92	0.87	0.87	0.87
N	630	495	210	179	281
$H_0 : \alpha_k = \alpha_{k'}$	$\alpha_0 = \alpha_2$		$\alpha_0 = \alpha_2$	$\alpha_0 = \alpha_1$	$\alpha_1 = \alpha_2$
F test:	0.34		23.38	1.14	9.48
p value	0.56		0.00	0.29	0.00
	$b^{k=0} = (2\alpha + \frac{7}{2}(1-\alpha)) + \alpha 100$ $b^{k=2} = (2\alpha + \frac{17}{5}(1-\alpha)) + \alpha 100 -$ $(300 + \frac{18}{5}s_i)(1-\alpha))\mathbb{1}_{[}s_i > \frac{500}{6}]$		$b^k = (2\alpha + 4(1-\alpha)) + \alpha 100$		

Notes: Robust standard errors clustered by individual subjects are reported in parentheses. *, **, and *** represent 10%, 5%, and 1% significance level, respectively.

There is substantial evidence of naive bidding in all treatments: an estimated parameter α is statistically significant at an usual significance level in all treatments of the English and second-price auctions. On the other hand, we observe the notable variations of $\widehat{\alpha}$ across insider treatments of the English auction. For the case of first drop-out prices (where the equilibrium strategies are all common across insider treatments), $\widehat{\alpha}$ decreases in the number of insiders: $\widehat{\alpha} = 0.66$ when $k = 0$; $\widehat{\alpha} = 0.53$ when $k = 1$; and $\widehat{\alpha} = 0.18$ when $k = 2$. The reduction of the degree of naive bidding is statistically significant relative to the treatment with two insiders. A similar pattern is established for the second drop-out prices. Comparing $\widehat{\alpha}$'s between the treatment with

$k = 0$ and that with $k = 1$ when the first drop-out bidder is an outsider, we again find a significant drop in the degree of naive bidding: $\widehat{\alpha} = 0.70$ when $k = 0$ and $\widehat{\alpha} = 0.32$ when $k = 1$ and the first drop-out bidder is an outsider. We do not find any statistical difference between two estimates in the treatment with $k = 1$ and 2 where the first drop-out bidder is an insider. In the second-price auction treatments, we do not find similar monotonic patterns of $\widehat{\alpha}$: $\widehat{\alpha} = 0.49$ when $k = 0$; $\widehat{\alpha} = 0.73$ when $k = 1$; and $\widehat{\alpha} = 0.40$ when $k = 2$.

The declining pattern of the degree of naive bidding in the English auction is quite intriguing. We conjecture that the presence of insiders–who have informational advantage–makes the outsider more wary about the information asymmetry and thus creates a motivation for the outsider to hedge against the asymmetry. This may work toward the correction of naive bidding and thus of the winner's curse in our setup. We summarize bidding behaviour in the experiment as follows.

Result 3 (bidding behaviour) (i) *There is evidence of naive bidding in both the second-price auction and the English auction, that is, overbidding relative to the Bayesian Nash equilibrium.* (ii) *The degree of naive bidding declines significantly in the increase of the number of insiders in the English auction.*

4.5 Conclusion

In this chapter we have proposed a model of interdependent value auctions with ex ante information asymmetry and examined key predictions of the model via a laboratory experiment. We study two standard auction formats–the second-price (sealed-bid) auction and the English auction. In each auction we allow any composition of insiders, who are perfectly informed of their value, and outsiders, who are only informed about the private component of their value. The information asymmetry between insiders and outsiders gives rise to potential mismatch of bidding strategies between them.

Our model is distinct from the existing auction literature with insider information in a couple of important respects (Engelbrecht-Wiggans et al. 1983, Hendricks & Porter 1988, Hendricks et al. 1994). First, unlike the literature where common value is typically assumed, we study the effects of insider information in standard auctions with interdependent value structure–the second-price (sealed-bid) auction and the English auction. The structure of interdependent value enables us to investigate the ramification of insider information on revenues as well as efficiency. We predict that the English auction has an efficient equilibrium, while the second-price auction is unable to avoid inefficiency caused by the presence of insider. Second, our theory is general in that we provide an equilibrium characterization for any number of insiders, in contrast to the common practice of the literature where only one insider is introduced. This yields novel insights to the literature. Most importantly, the increase in insider information by turning an outsider to an insider has a positive impact on the seller's revenues. This is reminiscent of the celebrated Linkage principle of Milgrom & Weber (1982*a*).

The experimental evidence supports the theoretical predictions on efficiency and revenues. We observe that subjects achieved an efficient allocation more frequently in the English auction than in the second-price auction when insider information is present. This is consistent with the theory that the English auction has an efficient equilibrium, whereas the second-price auction not. Controlling for realized signals, average revenues of the English auction increase in the number of insiders. Although we do not have revenue predictions for the second-price auction, we find similar patterns that revenues of the second-price auction increase in the number of insiders as well. Despite the compliance of experimental data to theory, there is substantial evidence of naive bidding as typical in the experimental auction literature (see Kagel & Levin (2011, 1995). Intriguingly, we find that the degree of naive bidding declines in the number of insiders in the English auction. We conjecture that the more insiders are present, the more wary outsiders are in their bidding behaviour, which may operate in the direction of reducing naive bidding and, as a result, the winner's curse. This is something to ponder over in developing an alternative, behavioural explanation of subjects' bidding behaviour.

4.6 Appendix

4.6.1 Proofs

Proof of Theorem 4.2.1: Suppose that bidder 1 is an insider and employs the undominated strategy of bidding $v_1(s)$ for each s. Given this, the optimal bid $b_2(s_2)$ of bidder 2 as an outsider must satisfy;

$$\begin{aligned} b_2(s_2) &< v_1(0, s_2) \text{ if } v_1(0, s_2) > v_2(0, s_2) \\ &= v_2(\alpha, s_2) \text{ if } v_1(\alpha, s_2) = v_2(\alpha, s_2) \text{ for some } \alpha \\ &> v_1(1, s_2) \text{ if } v_1(1, s_2) < v_2(1, s_2). \end{aligned} \tag{4.5}$$

Assume first that $v_1(0, s_2) > v_2(0, s_2)$, which implies by the single crossing property that $v_1(s_1, s_2) > v_2(s_1, s_2)$ for every $s_1 \in [0, 1]$, so it is efficient for bidder 1 to obtain the object regardless of s_1. Since bidder 1 bids $v_1(s_1, s_2)$, bidder 2 would incur a loss by winning. Bidder 2 could avoid this loss by bidding any $b < v_1(0, s_2) \leq v_1(s_1, s_2)$ and losing. A similar argument will establish that an optimal bid must be at least $v_2(1, s_2)$

if $v_1(1, s_2) < v_2(1, s_2)$, which will lead to the efficient allocation. Lastly, assume that $v_1(\alpha, s_2) = v_2(\alpha, s_2)$ for some α. Note that such α is unique due to the single crossing. Bidder 2's optimal bid b has to lie in the interval $[v_1(0, s_2), v_1(1, s_2)]$ so that there exists $\phi_1(b, s_2) \in [0, 1]$ such that $v_1(\phi_1(b, s_2), s_2) = b$. Letting $F_{S_1|S_2}(\cdot|s_2)$ denote the distribution of s_1 conditional on s_2, the expected payoff of bidder 2 with s_2 is

$$\int_0^{\phi_1(b,s_2)} (v_2(s_1, s_2) - v_1(s_1, s_2))dF_{S_1|S_2}s_1s_2.$$

The integrand is positive if and only if $s_1 < \alpha$, and thus the expression is maximized by setting $b = v_1(\alpha, s_2) = v_2(\alpha, s_2)$. Hence, bidder 2 wins if and only if $s_1 < \alpha$ or $v_1(s_1, s_2) < v_2(s_1, s_2)$ due to the single crossing, which means the resulting allocation is efficient. □

Proof of Theorem 4.2.2: Suppose to the contrary that there exists an efficient equilibrium of the second-price auction. For a given bidder i, let us define $E_i := \{s \in [0, 1]^n \mid v_i(s) \geq v_j(s) \text{ for all } j \neq i\}$, that is the set of signals for which bidder i wins the object at the efficient equilibrium. Due to the assumption that insiders obtain the good with some positive probability less than one, there must exist an outsider i, an insider j, and a signal profile s in the interior such that $v_i(s) = v_j(s) > \max_{k\neq i,j} v_k(s)$ (or $h_i(s_i) = h_j(s_j) > \max_{k\neq i,j} h_k(s_k)$). Fix any such profile s and let $E_{ij}(s_i) := \{s'|s'_i = s_i \text{ and } s' \in E_i \cap E_j\}$. Then, we can find another profile $\tilde{s} \in E_{ij}(s_i)$ such that $\tilde{s}_i = s_i$, $\tilde{s}_j = s_j$, and $\tilde{s}_k < s_k, \forall k \neq i, j$.

Now, given the efficient allocation and value bidding of bidder j, the bid $b_i(s_i)$ of bidder i with s_i has to satisfy

$$\max_{\{s'\in E_i|s'_i=s_i\}} v_j(s') \leq b_i(s_i) \leq \min_{\{s'\in E_j|s'_i=s_i\}} v_j(s'). \tag{4.6}$$

If the first inequality were violated, then bidder i with signal s_i would lose to bidder j when the former has a higher value. If the second inequality were violated, then bidder j would lose to bidder i with signal s'_i when the former has a higher value. Now, from (4.6)

$$\max_{s'\in E_{ij}(s_i)} v_j(s') \leq \max_{\{s'\in E_i|s'_i=s_i\}} v_j(s') \leq b_i(s_i) \leq \min_{\{s'\in E_j|s'_i=s_i\}} v_j(s') \leq \min_{s'\in E_{ij}(s_i)} v_j(s').$$

so $v_j(\cdot)$ has to be constant on $E_{ij}(s_i)$. This implies that for some constant k, $v_j(s') = k, \forall s' \in E_{ij}(s_i)$, which in turn implies that $v_i(s') = k, \forall s' \in E_{ij}(s_i)$ since $v_i(s') = v_j(s')$ for any $s' \in E_{ij}(s_i)$. Thus, for any $s' \in E_{ij}(s_i)$, we must have $h_i(s_i) = h_i(s'_i) = k - g(s')$, which cannot be true since, given our assumption, we have $g(s) > g(\tilde{s})$ even though

$s, \tilde{s} \in E_{ij}(s_i)$. $\square$

From now, we provide the proofs of Theorem 4.2.4 and 4.2.5 in Section 4.2.3. To simplify notation, we will let $s(p) := (s_{O\setminus A}, s_A(p; p_{N\setminus A}), s_{I\setminus A}(p; p_{N\setminus A}))$ and $s_i(p) = s_i(p; p_{N\setminus A})$ for $i \in I \cup A$, by omitting the price history $p_{N\setminus A}$. We first establish a couple of preliminary results in Lemma 4.6.1 and 4.6.3. To do so, let $\overline{v} = \max_{i\in N} \max_{s\in[0,1]^n} v_i(s)$. Given the assumption that $\lim_{s_i\to\infty} v_i(s_i, s_{-i}) \to \infty$ for any $s_{-i} \in [0,1]^{n-1}$, we can find some $\overline{s}_i$ for each i such that $v_i(\overline{s}_i, s_{-i}) \geq \overline{v}$ for any $s_{-i} \in [0,1]^{n-1}$.

Lemma 4.6.1. *For any $s_{O\setminus A}$ and $p_{N\setminus A}$, there exists a solution $(s_A, s_{I\setminus A}) : [\max_{i\in N\setminus A} p_i, \overline{v}] \to \times_{i\in I\cup A}[0, \overline{s}_i]$ of (4.3) such that for each $i \in A$, $s_i(\cdot)$ is strictly increasing.*

Proof. Let $v'_{A\cdot B}(s)$ denote a $|A| \times |B|$ matrix with its ij element being $\frac{\partial v_i}{\partial s_j}(s)$ for $i \in A$ and $j \in B$. From now, we write $v'_{A\cdot B}$ for convenience. Let 0_A and 1_A denote column vectors of 0's and 1's with dimension $|A|$, respectively. The following result from Krishna (2003) will be useful for the subsequent proof:

Lemma 4.6.2. *Suppose $A = (a_{ij})$ is an $m \times m$ matrix that satisfies the dominant average condition:*

$$\frac{1}{m}\sum_{k=1}^{m} a_{kj} > \quad a_{ij}, \forall i \neq j \quad \text{and} \quad \sum_{k=1}^{m} a_{kj} > 0, \forall j. \tag{4.7}$$

Then, A is invertible. Also, there exists a unique $x \gg 0$ such that $Ax = 1$, where 1 is a column vector of m 1's.

To obtain a solution to (4.3) recursively, suppose that the set of active bidders is A and the unique solution of (4.3) exists up to price $\overline{p} = \max_{k\in N\setminus A} p_k$. Let $(\underline{s}_A, \underline{s}_{I\setminus A})$ denote this solution at $\overline{p}$. We extend the solution beyond $\overline{p}$ to all $p \in [\overline{p}, \overline{v}]$. To do so, differentiate both sides of (4.3) with p to get the following differential equation:

$$\begin{pmatrix} v'_{A\cdot A} & v'_{A\cdot I\setminus A} \\ v'_{I\setminus A\cdot A} & v'_{I\setminus A\cdot I\setminus A} \end{pmatrix} \begin{pmatrix} s'_A(p) \\ s'_{I\setminus A}(p) \end{pmatrix} = \begin{pmatrix} 1_A \\ 0_{I\setminus A} \end{pmatrix} \tag{4.8}$$
$$(s_A(\overline{p}), s_{I\setminus A}(\overline{p})) = (\underline{s}_A, \underline{s}_{I\setminus A}).$$

Note that the first matrix in the LHS can be written as $v'_{I\cup A\cdot I\cup A}$. Assume for the moment that $v'_{N\cdot N}$ is invertible so its principal minors $v'_{I\cup A\cdot I\cup A}$ and $v'_{I\setminus A\cdot I\setminus A}$ are invertible too. Then, by Peano's theorem, a unique solution of (4.8) exists since the value functions are twice continuously differentiable. We now show that $v'_{N\cdot N}$ is invertible and also $s_A(p)' \gg 0$. To do so, let us rewrite the last $|I\setminus A|$ lines of (4.8) as

$s'_{I\backslash A} = -(v'_{I\backslash A \cdot I\backslash A})^{-1} v'_{I\backslash A \cdot A} s'_A$. Substitute this into the first $|A|$ lines of (4.8) to obtain $V s'_A = 1_A$ after rearrangement, where

$$V := v'_{A\cdot A} - v'_{A\cdot I\backslash A}(v'_{I\backslash A\cdot I\backslash A})^{-1} v'_{I\backslash A\cdot A}.$$

If V satisfies the dominant average condition for any A, then, with $A = N$, $V = v'_{N\cdot N}$ is invertible by Lemma 4.6.2. Also, by Lemma 4.6.2, $s'_A(p) \gg 0$.

To prove that V satisfies the dominant average condition, let $g'_k = \frac{\partial g}{\partial s_k}$ and $g'_A = (g'_k)_{k\in A}$, where g'_A is considered as a column vector. Let D_A denote the diagonal matrix whose diagonal entry is $h'_k = \frac{dh_k}{ds_k}$ for $k \in A$. Then, for any $A, B \subset N$,

$$v'_{A\cdot B} = \begin{cases} D_A + 1_A (g'_A)^t & \text{if } A = B \\ 1_A (g'_B)^t & \text{if } A \cap B = \emptyset, \end{cases}$$

where $(\cdot)^t$ denotes the transpose of a matrix. Using this, we can rewrite V as

$$\begin{aligned} V &= D_A + 1_A(g'_A)^t - 1_A(g'_{I\backslash A})^t \left(D_{I\backslash A} + 1_{I\backslash A}(g'_{I\backslash A})^t\right)^{-1} 1_{I\backslash A}(g'_A)^t \\ &= D_A + (1-x)1_A(g'_A)^t, \end{aligned} \tag{4.9}$$

where $x = (g'_{I\backslash A})^t \left(D_{I\backslash A} + 1_{I\backslash A}(g'_{I\backslash A})^t\right)^{-1} 1_{I\backslash A}$. Since all the entries in any given column of the matrix $1_A(g'_A)^t$ are identical and the diagonal entries of D_A are positive, the first inequality of (4.7) is easily verified. The proof will be complete if the second inequality of (4.7) is shown to hold, for which it suffices to show $x < 1$:

$$\begin{aligned} x &= (g'_{I\backslash A})^t \left(D_{I\backslash A} + 1_{I\backslash A}(g'_{I\backslash A})^t\right)^{-1} 1_{I\backslash A} \\ &= (g'_{I\backslash A})^t \left(D^{-1}_{I\backslash A} - \left(\frac{1}{1 + (g'_{I\backslash A})^t D^{-1}_{I\backslash A} 1_{I\backslash A}}\right) D^{-1}_{I\backslash A} 1_{I\backslash A} (g'_{I\backslash A})^t D^{-1}_{I\backslash A}\right) 1_{I\backslash A} \\ &= (g'_{I\backslash A})^t D^{-1}_{I\backslash A} 1_{I\backslash A} - \frac{\left((g'_{I\backslash A})^t D^{-1}_{I\backslash A} 1_{I\backslash A}\right)^2}{1 + (g'_{I\backslash A})^t D^{-1}_{I\backslash A} 1_{I\backslash A}} \\ &= \frac{(g'_{I\backslash A})^t D^{-1}_{I\backslash A} 1_{I\backslash A}}{1 + (g'_{I\backslash A})^t D^{-1}_{I\backslash A} 1_{I\backslash A}} = \frac{\sum_{k\in I\backslash A} g'_k / h'_k}{1 + \sum_{k \in I\backslash A} g'_k/h'_k} < 1, \end{aligned}$$

where the second equality was derived using the formula for an inverse matrix,

$$\left(A + bc^t\right)^{-1} = A^{-1} - \left(\frac{1}{1 + c^t A^{-1} b}\right) A^{-1} b c^t A^{-1},$$

with $A = D_{I\backslash A}$, $b = 1_{I\backslash A}$, and $c = g'_{I\backslash A}$. $\square$

Given the break-even signals obtained in Lemma 4.6.1, we consider each outsider i's strategy of dropping out (staying in) at p if and only if $s_i < s_i(p)$ after any history $p_{N\backslash A}$. Together with the insiders' value bidding strategy, we refer to this strategy profile as β^*.

Lemma 4.6.3. *Given the strategy profile β^*, for any signal profile $s \in [0,1]^n$, (i) outsiders drop out in order of their values; (ii) for each outsider i, $p_i \leq v_i(s)$; and (iii) at any outsider i's drop-out price p_i, $s_j(p_i) \geq s_j$ for each insider j who is inactive at p_i.*

Proof. Consider two outsiders i and j with $p_i \leq p_j$. Then, at the price p_i at which bidder i drops out, we have

$$h_i(s_i) = h_i(s_i(p_i)) = p_i - g(s(p_i)) = h_j(s_j(p_i)) \leq h_j(s_j),$$

where the first equality and the inequality follow from the drop-out strategy of the outsiders i and j, respectively, while the second and third equalities from the break-even condition at p_i. This proves (i) since $h_i(s_i) \leq h_j(s_j)$ means $v_i(s) \leq v_j(s)$.

To prove (ii), suppose to the contrary that $p_i > v_i(s)$. Since $s_i(p_i) = s_i$, this means $h_i(s_i) + g(s(p_i)) = p_i > h_i(s_i) + g(s)$, so $g(s(p_i)) > g(s)$. Given this, for each insider $j \in I$ (whether active or not), we must have

$$h_j(s_j) - h_j(s_j(p_i)) \geq g(s(p_i)) - g(s) > 0, \tag{4.10}$$

where the first inequality holds since the break-even condition implies that for an inactive insider j, $h_j(s_j) + g(s) = v_j(s) = p_j = v_j(s(p_i)) = h_j(s_j(p_i)) + g(s(p_i))$ while for an active insider j, $h_j(s_j) + g(s) = v_j(s) \geq p_i = v_j(s(p_i)) = h_j(s_j(p_i)) + g(s(p_i))$. The inequality (4.10) implies $s_j > s_j(p_i)$ for each insider j. Also, for each active outsider $j \in O \cap A$, we have $s_j \geq s_j(p_i)$. Thus, $s \geq s(p_i)$ so $v_i(s) \geq v_i(s(p_i)) = p_i$, a contradiction.

To prove (iii), note first that due to (ii), we have $h_i(s_i) + g(s) = v_i(s) \geq p_i = h_i(s_i(p_i)) + g(s(p_i)) = h_i(s_i) + g(s(p_i))$ since $s_i(p_i) = s_i$. This inequality means $g(s) \geq g(s(p_i))$. If an insider j is inactive at p_i, then we must have $v_j(s) = p_j = v_j(s(p_i))$ or $h_j(s_j(p_i)) - h_j(s_j) = g(s) - g(s(p_i)) \geq 0$, which yields $s_j(p_i) \geq s_j$. $\square$

Proof of Theorem 4.2.4: We first prove that the strategy profile β^*, if followed by all bidders, leads to the efficient allocation, and then show that it constitutes an ex-post equilibrium.

Given that outsiders drop out in order of their values (according to (i) of Lemma 4.6.3), the efficiency result will follow if an outsider i with the highest value among outsiders drops out before (after) an insider j with the highest value among insiders if and only if $v_i(s) < (>) v_j(s)$. In case $v_i(s) < v_j(s)$, the outsider i dropping out at some $p_i \leq v_i(s)$ (from (ii) of Lemma 4.6.3) means that the insider j is a winner since $p_i \leq v_i(s) < v_j(s)$ is lower than the insider j's drop-out price $v_j(s)$. Assume now that $v_i(s) > v_j(s)$ and suppose to the contrary that the outsider i drops out at some price $p_i < v_j(s)$ at which only insiders, including j, are active.[17] Then, the break-even condition at p_i implies $h_i(s_i) = p_i - g(s(p_i)) = h_k(s_k(p_i))$ for each $k \in I \cap A$. Since $h_i(s_i) > h_k(s_k)$ for all those k, this means $h_k(s_k(p_i)) > h_k(s_k)$ or $s_k(p_i) > s_k$. Thus, due to (iii) of Lemma 4.6.3, we have $s(p_i) = (s_O, s_{I\cap A}(p_i), s_{I\setminus A}(p_i)) \geq s$ with $s_k(p_i) > s_k$, which implies $p_i = v_k(s(p_i)) > v_k(s)$ for all $k \in I \cap A$. This contradicts with the value bidding strategy of insiders.

To show that the strategy profile β^* constitutes an ex-post equilibrium, let us focus on an arbitrary outsider i. If i has the highest value and follows the equilibrium strategy to be a winner, then his payoff is $v_i(s) - \max_{k\neq i} p_k \geq v_i(s) - \max_{k\neq i} v_k(s) \geq 0$.[18] So any nontrivial deviation by i cannot be profitable since it will only result in losing and earning zero payoff. Suppose now that there is some j with $v_j(s) > v_i(s)$. If j is an insider, then any nontrivial deviation by i to become a winner would make him pay at least $v_j(s)$, i.e. more than his value. Let us thus focus on the case j is an outsider with the highest value. Any nontrivial deviation by i would require him to wait beyond some price p such that $s_i(p) = s_i$, and then becoming a winner after j drops out last at some $p_j > p$.[19] Then, we must have $s_i(p_j) > s_i$ and $s_j(p_j) = s_j$. Combining this with $s_I(p_j) \geq s_I$ (from (iii) of Lemma 4.6.3), we have $s(p_j) = (s_{O\setminus\{i\}}, s_i(p_j), s_I(p_j)) \geq s$ with $s_i(p_j) > s_i$, so $v_i(s(p_j)) = p_j > v_i(s)$, implying the deviation would incur a loss to i. □

Proof of Theorem 4.2.5: Throughout the proof, for any variable x in E, we let x' denote its counterpart in E'. For instance, p'_k denotes a drop-out price of bidder k in E'. Let $O = \{1, 2, \cdots, l\}$ and thus $I = \{l+1, \cdots, n\}$, and assume that $v_1(s) \leq v_2(s) \leq \cdots \leq v_l(s)$, without loss of generality. Then, $O' = O\setminus\{i\}$ and $I' = I \cup \{i\}$.

First, according to (ii) of Lemma 4.6.3, the switched insider i drops out at a (weakly) higher price in E' than in E. The proof will be complete if we show that all other outsiders drop out at (weakly) higher prices in E' as well. Now suppose for a con-

[17]This holds since bidder i is the last to drop out among outsiders, according to (i) of Lemma 4.6.3.

[18]The first inequality holds since each insider drops out at his value and each outsider drops out below his value according to (ii) of Lemma 4.6.3.

[19]An argument similar to that in the proof of the efficiency can be used to show that since j has the highest value, j is the one to drop out last (except for i) even under i's deviation.

tradiction that some outsider drops out at a lower price in E' than in E. Note that for all outsider $k < i$, we have $p'_k = p_k$ since the history of drop-out prices is the same across E and E' until p_i is reached. Using this and our assumption, let us define $j = \min\{k \,|\, p'_k < p_k, \text{ and } i < k \leq l\}$. Let us first make a few of observations: (i) a signal s_k for each $k < j$ with $k \neq i$ has been revealed in both E and E' when the price clock reaches p'_j, since $p_k \leq p'_k \leq p'_j$ for all such k[20]; (ii) $s'_j(p'_j) = s_j = s_j(p_j) > s_j(p'_j)$ since $p_j > p'_j$; and (iii) $s'_i(p'_j) \geq s_i$. To see (iii), it follows from (iii) of Lemma 4.6.3 if i is inactive at p'_j in E'. If i is active at p'_j, then the monotonicity of $s'_i(\cdot)$ implies $s'_i(p'_j) \geq s'_i(p_i) = s_i(p_i) = s_i$ since $p'_j \geq p_i$.[21] We next show that

$$s'_k(p'_j) \geq s_k(p'_j) \text{ for all } k \in \{j+1, \cdots, n\}, \tag{4.11}$$

which, given (i), (ii), and (iii) above, will imply that $s'(p'_j) \geq s(p'_j)$[22] with $s'_j(p'_j) > s_j(p'_j)$, so $p'_j = v_j(s'(p'_j)) > v_j(s(p'_j)) = p'_j$, yielding the desired contradiction. To prove (4.11), observe first that the break-even conditions at price p'_j in E and E' yield

$$g(s(p'_j)) = p'_j - h_j(s_j(p'_j)) > p'_j - h_j(s'_j(p'_j)) = g(s'(p'_j)),$$

where the inequality holds due to (ii) above. We then prove (4.11) by considering two cases depending on whether or not $k \in \{j+1, \cdots, n\}$ is active at p'_j in E'. Since each outsider $k \in \{j+1, \cdots, l\}$ is active at p'_j in E', an inactive bidder $k \in \{j+1, \cdots, n\}$ must be an insider. For such k, we obtain (4.11) since the break-even conditions at price p'_j in E and E' yield

$$h_k(s'_k(p'_j)) = p_k - g(s'(p'_j)) > p_k - g(s(p'_j)) = h_k(s_k(p'_j)), \tag{4.12}$$

where the inequality follows from (4.12). Turning to the case in which bidder $k \in \{j+1, \cdots, n\}$ is active at p'_j in E', he must be active at p'_j in E as well. The reason is that if k is an outsider, then $p'_j < p_j$ and he drops out no sooner than j in E (due to (i) of Lemma 4.6.3) while if k is an insider, he drops out at the same price (i.e. his value) in E and E'. Thus, we obtain (4.11) since the break-even conditions at p'_j in E and E' yield

$$h_k(s'_k(p'_j)) = p'_j - g(s'(p'_j)) > p'_j - g(s(p'_j)) = h_k(s_k(p'_j)),$$

[20] The second inequality here holds since outsiders drop out in order of their values in E'.

[21] The equality $s'_i(p_i) = s_i(p_i)$ follows from the fact that the price history is the same across E and E' until p_i is reached.

[22] Note that the i-th component of $s(p_j)$ is equal to s_i so the inequality follows from (iii).

where the inequality follows again from (4.12). □

4.6.2 Experimental analysis

Table 4.11: Regression analysis of revenues more flexible specifications

Variables	English	SBSP
k = 1	10.847***	10.861***
	(1.986)	(2.468)
k = 2	18.853***	10.435***
	(1.74)	(2.214)
s(1)	0.975***	0.288
	(.227)	(.304)
s(2)	2.037***	1.710***
	(.153)	(.197)
s(3)	0.466**	1.075***
	(.226)	(.301)
$s(1)^2$	0.006***	0.009***
	(.002)	(.003)
$s(2)^2$	0.001	0.003
	(.002)	(.002)
$s(3)^2$	0.002	-0.004
	(.002)	(.002)
s(1) x s(2)	-0.007*	-0.014***
	(.003)	(.004)
s(1) x s(3)	0	0.009**
	(.003)	(.004)
constant	2.36	15.850*
	(6.665)	(9.04)
# of obs.	1125	975
R2	0.908	0.839
p-value H0: (k=0) = (k=1)	0.000	0.000
p-value H0: (k=1) = (k=2)	0.000	0.834

Notes: Standard errors are reported in parentheses. *, **, and *** represent 10%,5%, and 1% significance level, respectively. $s_{(1)} = min[s]$, $s_{(3)} = max[s]$, $s_{(2)} = med[s]$

Table 4.12: Regression analysis of revenues more flexible specifications

Variables	English	SBSP
k = 1	15.665**	17.633*
	(7.936)	(9.763)
k = 2	29.069*	24.29
	(15.381)	(18.879)
s(1)	0.784***	0.626***
	(.048)	(.066)
s(2)	2.053***	1.716***
	(.043)	(.059)
s(3)	0.707***	0.622***
	(.051)	(.066)
$\mathbb{1}_{[s(1)=s(Insider)]}$	-10.971	-16.276*
	(7.892)	(9.833)
$\mathbb{1}_{[s(2)=s(Insider)]}$	-5.224	-19.698*
	(8.303)	(10.228)
$\mathbb{1}_{[s(3)=s(Insider)]}$	-23.238**	-31.405***
	(9.224)	(11.549)
$\mathbb{1}_{[s(1)=s(Insider)]} \times s(1)$	0.245***	0.261***
	(.067)	(.092)
$\mathbb{1}_{[s(2)=s(Insider)]} \times s(2)$	0.082	0.332***
	(.058)	(.075)
$\mathbb{1}_{[s(3)=s(Insider)]} \times s(3)$	0.181***	0.310***
	(.066)	(.09)
constant	3.565	38.074***
	(4.217)	(5.168)
# of obs.	1125	975
R2	0.91	0.844
p-value H0: (k=0) = (k=1)	0.049	0.071
p-value H0: (k=1) = (k=2)	0.084	0.484

Notes: Standard errors are reported in parentheses. *, **, and *** represent 10%,5%, and 1% significance level, respectively. $s_{(1)} = min[s]$, $s_{(3)} = max[s]$, $s_{(2)} = med[s]$

4.6.3 Maximum Likelihood approach

The ML approach deals with the censoring existing in the observed data. It distinguishes between the observed drop–out bid $d_{s,ir}$ and the reservation bid $p_{s,ir}$. For easiness of exposition denote, for all $i \in N$, $g_1(i,r)$ as the individual who first drops out from i's group in round r, and the corresponding group as $g(i,r)$, let $G_1(r)$ be the collection of all first bidders at round r. Therefore, for all $i \in G_1(r)$ we know $p_{1,ir} = d_{1,ir}$ but for the other two active bidders $j \in g(i,r) \setminus g_1(i,r)$ we only know that $p_{1,jr} > p_{1,g_1(j,r)r}$ which means that we observe a right–censored variable of their true drop–out price. For the second stage, we know that for bidder $g_2(i,r)$, who second drops out, his/her reservation bid is $p_{2,g_2(i,r)r} = d_{2,g_2(i,r)r}$, but for the remaining bidder j we only know $d_{2,jr} > d_{2,g_2(i,r)r}$, which again implies a right–censored variable.

Consider $\epsilon^k_{s,ir} = p^k_{s,ir} - \Gamma^k_{s,ir}$ where $\Gamma^k_{s,ir}$ follows from the right hand side of the regression equation. We know that $d_{s,g_1(i,r)r} = p_{s,ir}$ if and only if $i \in g_s(i,r)$. If we define $e^k_{s,ir} = d^k_{s,g_1(i,r)r} - \Gamma^k_{s,ir}$ then $d_{s,g_1(i,r)r} = p_{s,ir}$ if and only if $e^k_{s,ir} = \epsilon^k_{s,ir}$.

Denoting $\theta^k_s = \left(\alpha_{\mathbf{s}}{}^k, \beta_{\mathbf{s}}{}^k, \delta_{\mathbf{s}}{}^k, \{\sigma_{s,i} : i : 1 \to N_k\}\right)$ and $D_s \equiv \left(d_{s,g_s(i,r)}\right)_{\forall i \in N_k, r \in R}$ the information on drop–out prices from the experiment, the density function associated to the first bidding function $p_{1,ir}$ is given by

$$f_{p_{1,ir}}(b \mid \cdot) = f(p_{1,ir} = d_{1,g_1 i,rr} \mid \cdot)^{\mathbf{1}_{[i \in g_1(i,r)]}} (1 - F(p_{1,ir} \leq d_{1,g_1 i,rr} \mid \cdot))^{\mathbf{1}_{[i \notin g_1(i,r)]}},$$

therefore the ML function is

$$L^k_1\left(\theta^k_1; D_1\right) = \prod_{r \in R} \prod_{i \in N_k} \left[\frac{1}{\sigma_i} \phi\left(\frac{e^k_{1,ir}}{\sigma_i}\right)\right]^{\mathbf{1}_{[i \in g_1(i,r)]}} \left[1 - \Phi\left(\frac{e^k_{1,ir}}{\sigma_i}\right)\right]^{\mathbf{1}_{[i \notin g_1(i,r)]}} \tag{4.13}$$

On the other hand, the ML associated to second bidders is

$$L^k_2\left(\theta^k_2; D_2\right) = \prod_{r \in R} \prod_{i \in N_k \setminus G_1(r)} \left[\frac{1}{\sigma_i} \phi\left(\frac{e^k_{2,ir}}{\sigma_i}\right)\right]^{\mathbf{1}_{[i \in g_2(i,r)]}} \left[1 - \Phi\left(\frac{e^k_{2,ir}}{\sigma_i}\right)\right]^{\mathbf{1}_{[i \notin g_2(i,r)]}} \tag{4.14}$$

Notice that the specification corresponds to a Partial ML estimator. From Wooldridge (2003) we know that, once correcting the variance matrix for within–subjects dependence, the pooled partial MLE analysis is consistent and asymptotically normal.

4.6.4 Supplementary Material

Equilibrium of Second-Price Auction with $I = \{2, 3\}$: We aim to find bidder 1's bid that is a best response to the value bidding of the two insiders, bidder 2 and

3. By symmetry, it suffices to focus on the case in which $s_2 \geq s_3$ so $v_2(s) \geq v_3(s)$, meaning bidder 2 bids higher than bidder 3. Then, by bidding b, bidder 1 wins the object to obtain a payoff equal to $v_1(s) - v_2(s) = (a-1)(s_1 - s_2)$ when $b \geq v_2(s)$ and $s_2 \geq s_3$, which can be rewritten as $s_3 \leq \min\{b - as_2 - s_1, s_2\}$. Given this and the uniform distribution of signals, bidder 1's payoff from biding b is given as

$$\pi(b; s_1) = \int_0^{\min\{1, \frac{b-s_1}{a}\}} (a-1)(s_1 - s_2) \min\{b - as_2 - s_1, s_2\} ds_2.$$

As one can check, this expression is maximized by setting $b = B_1(s_1)$ with $B_1(s_1)$ defined in (4.2).

Equilibrium of Second-Price Auction with $I = \{3\}$: Let $B : [0,1] \to \mathbb{R}_+$ denote a symmetric, non-decreasing bidding strategy for the two outsiders. We first prove Proposition 4.2.3 to obtain a partial characterization of monotone equilibrium bidding strategy for general value distribution:

Proof of Proposition 4.2.3. To first show that $B(s) \geq (a+1)s$ for all $(0,1]$, suppose for a contradiction that bidder 1, for instance, with some signal $\hat{s} \in (0,1]$ is bidding $B(\hat{s}) < (a+1)\hat{s}$. We consider this bidder's payoff at the margin, i.e. when his bid is tied with that of either bidder 2 or bidder 3 as the highest bid. First, being tied with bidder 2 means that $s_2 = \hat{s}$, in which case bidder 1's (ex-post) payoff is equal to $v_1(s) - B(\hat{s}) > a\hat{s} + \hat{s} + s_3 - (a+1)\hat{s} = s_3 \geq 0$. Next, being tied with bidder 3 means that $v_3(s) = as_3 + \hat{s} + s_2 = B(\hat{s}) < (a+1)\hat{s}$, which implies that $as_3 < a\hat{s} - s_2$ so $s_3 < \hat{s}$. In this case, bidder 1's ex-post payoff is equal to $v_1(s) - v_3(s) = (a-1)(\hat{s} - s_3) > 0$. In sum, bidder 1 with $\hat{s}$ obtains a positive marginal payoff whether he is tied with bidder 2 or bidder 3. So, it is profitable to slight increase his bid from $B(\hat{s})$.

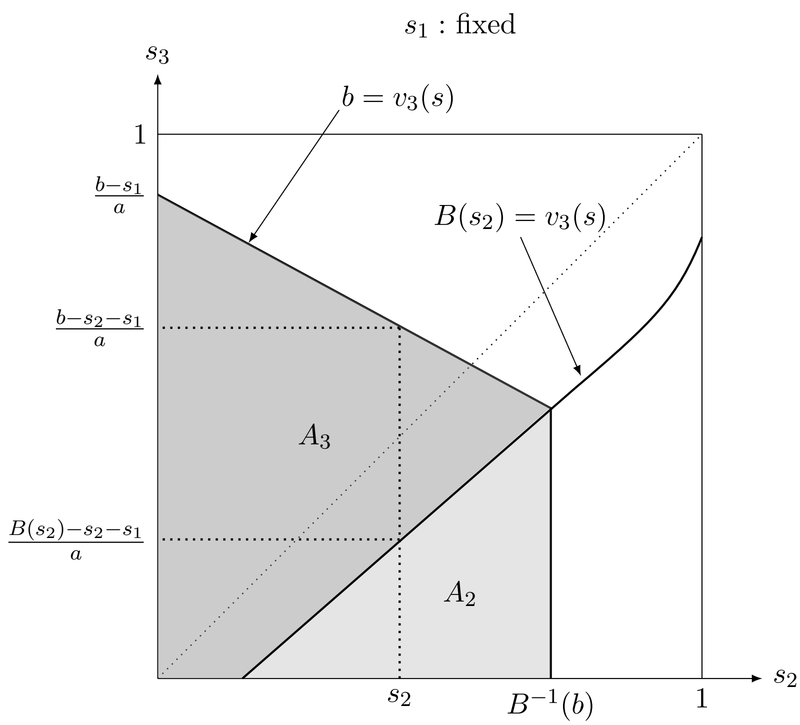
To next show that $B(s) \leq (a+2)s$ for all $s \in [0,1]$, let now suppose to the contrary that bidder 1, for instance, with some signal $\hat{s} \in [0,1]$ is bidding $B(\hat{s}) > (a+2)\hat{s}$. As above, we consider this bidder's payoff at the margin. In the case of tying with bidder 2, we have $s_2 = \hat{s}$ and $v_3(s) = as_3 + 2\hat{s} \leq B(\hat{s})$, which implies $s_3 \leq \frac{B(\hat{s}) - 2\hat{s}}{a}$. So bidder 1's (ex-post) payoff is

$$(a+1)\hat{s} + s_3 - B(\hat{s}) \leq (a+1)\hat{s} + \frac{B(\hat{s}) - 2\hat{s}}{a} - B(\hat{s}) = \frac{a-1}{a}[B(\hat{s}) - (a+2)\hat{s}] < 0$$

since $B(\hat{s}) > (a+2)\hat{s}$. In the case of the tying with bidder 3, we have $s_2 \leq \hat{s}$ and $v_3(s) = as_3 + \hat{s} + s_2 = B(\hat{s}) > (a+2)\hat{s}$, which implies $s_3 > \hat{s}$. Then, the payoff of bidder 1 with $\hat{s}$ is $v_1(s) - v_3(s) = (a-1)(\hat{s} - s_3) < 0$. In sum, bidder 1 with $\hat{s}$ obtains a negative marginal payoff whether he is tied with bidder 2 or bidder 3. So, it is profitable to slightly reduce his bid from $B(\hat{s})$.

To show (ii), note that by (i), we must have $B(1) \leq (a+2)$. Suppose now for a contradiction that $B(1) < a+2$. If an outsider i with signals $s_i = 1$ deviates to some bid greater than $B(1)$, then it will only increase his chance of winning against bidder 3, in which case his payoff increases by $v_i(1, s_{-i}) - v_3(1, s_{-i}) = (a-1)(1-s_3) > 0$. So the deviation is profitable. □

To go beyond the partial characterization in Proposition 4.2.3, let us consider the problem faced by bidder 1 with any fixed signal $s_1 \in [0,1]$. By bidding b, he wins if $s_2 \leq B^{-1}(b)$ and $v_3(s) \leq b$. Then, his payment is equal to $v_3(s)$ if $v_3(s) \geq B(s_2)$ and equal to $B(s_2)$ otherwise. In the former case (the darker gray area A_3 in the graph below), his (ex-post) payoff is $v_1(s) - v_3(s) = (a-1)(s_1 - s_3)$ while in the latter case (the lighter gray area A_2 in the graph below), his payoff is $v_1(s) - B(s_2) = as_1 + s_2 + s_3 - B(s_2)$.



Given this and the uniform distribution of signals, the expected payoff of bidder 1 with s_1 can be written as

$$\pi(b;s_1) = \int_0^{B^{-1}(b)} \left[\int_{\max\{\frac{B(s_2)-s_2-s_1}{a},0\}}^{\min\{\frac{b-s_2-s_1}{a},1\}} (a-1)(s_1-s_3)ds_3 \right] ds_2$$
$$+ \int_0^{B^{-1}(b)} \left[\int_0^{\max\{\frac{B(s_2)-s_2-s_1}{a},0\}} (as_1+s_2+s_3-B(s_2))ds_3 \right] ds_2.$$

The first (resp., second) integration corresponds to bidder 1's payoff in the area A_3 (resp,. A_2). Then, the requirement that this payoff be maximized by setting $b = B(s_1)$ gives rise to a differential equation with which we can solve for B. While we omit the detailed expression for the differential equation, it yields a linear solution below a threshold signal:

$$B(s_1) = \left(\frac{-7+7a+4a^2+\sqrt{1+14a-23a^2-8a^3+16a^4}}{2(-3+4a)} \right) s_1 \text{ for } s_1 \in [0,\overline{s}],$$

where the threshold $\overline{s}$ is the signal s_1 that solves $\frac{B(s_1)-s_1}{a} = 1$.[23] Unfortunately, an analytical form solution for B is unavailable beyond the range $[0,\overline{s}]$. Instead, we have used a numerical method to draw a graph of B that is given as follow:

[23] We note that for $a = 2$, $B(s_1) \simeq 3.44 s_1$ and $\overline{s} \simeq 0.82$

Experimental instructions

Instructions SPSB $k = 0$

Instructions

This is an experiment in the economics of decision-making. Research foundations have provided funds for conducting this research. Your earnings will depend partly on your decisions and partly on the decisions of the other participants in the experiment. If you follow the instructions and make careful decisions, you may earn a considerable amount of money.

At this point, check the name of the computer you are using as it appears on the top of the monitor. At the end of the experiment, you should use your computer name to claim your payments. At this time, you will receive £5 as a participation fee simply for showing up on time. In addition, you will receive £10 as an initial balance you will use in this experiment. Any positive or negative earnings incurred during the experiment will be added into this balance. Details of how you will make decisions will be provided below. During the experiment we will speak in terms of experimental tokens instead of pounds. Your payments will be calculated in terms of tokens and then exchanged at the end of the experiment into pounds at the following rate:

40 Tokens = 1 Pound

In this experiment, you will participate in 17 independent and identical (of the same form) auction rounds. In each round you will act as a bidder in an auction and compete for a single hypothetical object with other two participants in your group. Note that the first two rounds are practice rounds in which your earnings will not be counted for actual payoffs. The remaining 15 auction rounds are real and any positive or negative earnings will be counted for actual payoffs. If your balance during the experiment goes below zero, you will become inactive and be excluded for any remaining auction rounds and will receive only £5 participation fee at the end of the experiment.

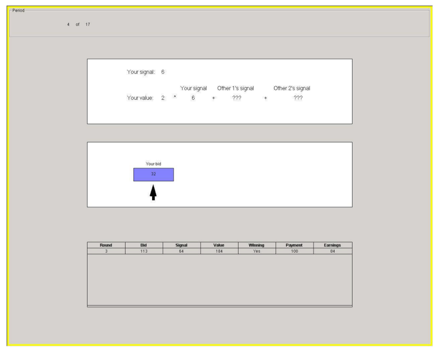
An auction round

Next, we will describe in detail the process that will be repeated in all 17 rounds. Each round starts by having the computer randomly form three-person groups. The groups formed in each round depend solely upon chance and are independent of the groups formed in any of the other rounds. That is, in any group each active participant is equally likely to be chosen for that group. In a case where any other participant was excluded due to its negative balance, there is a chance that you may become inactive in a particular round when you are matched with that participant who was excluded.

In the beginning of each round, each participant will be assigned a *signal* that will be randomly drawn from the set of integer numbers of tokens between 0 and 100 (numbers not including decimals). That is, any number from the set {0, 1, 2, …, 100} will be equally likely to be drawn. A signal you will be assigned in each round is independent of signals other participants will be assigned and is independent of a signal assigned to you in any of the other rounds. This will be done by the computer.

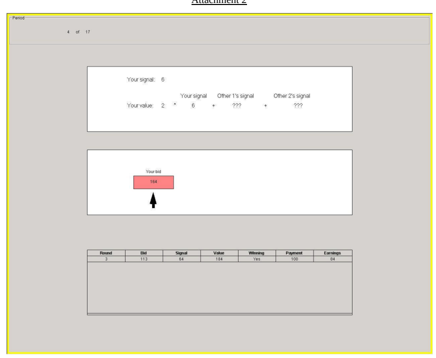
1

Attachment 1



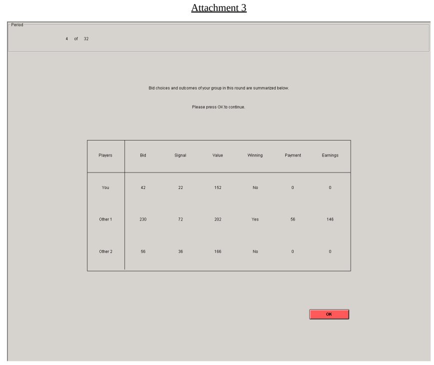
1

Attachment 2



2

Attachment 3



3

Figure 4.6: Attachments Instructions SPSB $k = 0$

Instructions SPSB $k = 1$

<u>Instructions</u>

This is an experiment in the economics of decision-making. Research foundations have provided funds for conducting this research. Your earnings will depend partly on your decisions and partly on the decisions of the other participants in the experiment. If you follow the instructions and make careful decisions, you may earn a considerable amount of money.

At this point, check the name of the computer you are using as it appears on the top of the monitor. At the end of the experiment, you should use your computer name to claim your payments. At this time, you will receive £5 as a participation fee simply for showing up on time. In addition, you will receive £10 as an initial balance you will use in this experiment. Any positive or negative earnings incurred during the experiment will be added into this balance. Details of how you will make decisions will be provided below. During the experiment we will speak in terms of experimental tokens instead of pounds. Your payments will be calculated in terms of tokens and then exchanged at the end of the experiment into pounds at the following rate:

40 Tokens = 1 Pound

In this experiment, you will participate in 17 independent and identical (of the same form) auction rounds. In each round you will act as a bidder in an auction and compete for a single hypothetical object with other two participants in your group. Note that the first two rounds are practice rounds in which your earnings will not be counted for actual payoffs. The remaining 15 auction rounds are real and any positive or negative earnings will be counted for actual payoffs. If your balance during the experiment goes below zero, you will become inactive and be excluded for any remaining auction rounds and will receive only £5 participation fee at the end of the experiment.

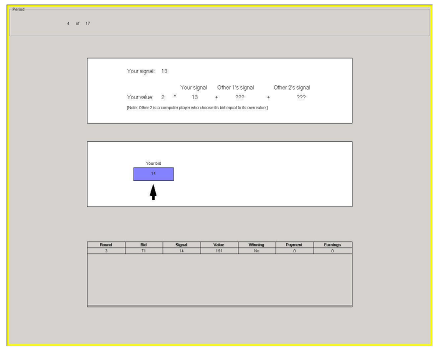
<u>An auction round</u>

Next, we will describe in detail the process that will be repeated in all 17 rounds. Each round starts by having the computer randomly form three-participant groups. In each group, one participant is played by the computer (called a computer participant), while the other two participants are played by persons. The groups formed in each round depend solely upon chance and are independent of the groups formed in any of the other rounds. That is, in any group each active person is equally likely to be chosen for that group. In a case where any other person was excluded due to its negative balance, there is a chance that you may become inactive in a particular round when you are matched with that person who was excluded.

In the beginning of each round, each participant will be assigned a *signal* that will be randomly drawn from the set of integer numbers of tokens between 0 and 100 (numbers not including decimals). That is, any number from the set {0, 1, 2, …, 100} will be equally likely to be drawn. A signal you will be assigned in each round is independent of signals other participants will be assigned and is independent of a signal assigned to you in any of the other rounds. This will be done by the computer.

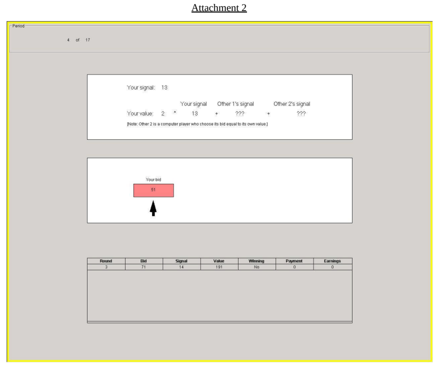
1

Attachment 1



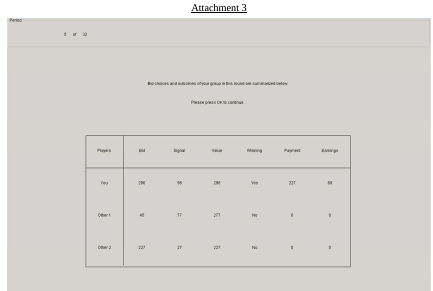
1

Attachment 2



2

Attachment 3



3

Figure 4.7: Attachments Instructions SPSB any $k \in \{1, 2\}$

Instructions SPSB $k = 2$

Instructions

This is an experiment in the economics of decision-making. Research foundations have provided funds for conducting this research. Your earnings will depend partly on your decisions and partly on the decisions of the other participants in the experiment. If you follow the instructions and make careful decisions, you may earn a considerable amount of money.

At this point, check the name of the computer you are using as it appears on the top of the monitor. At the end of the experiment, you should use your computer name to claim your payments. At this time, you will receive £5 as a participation fee simply for showing up on time. In addition, you will receive £10 as an initial balance you will use in this experiment. Any positive or negative earnings incurred during the experiment will be added into this balance. Details of how you will make decisions will be provided below. During the experiment we will speak in terms of experimental tokens instead of pounds. Your payments will be calculated in terms of tokens and then exchanged at the end of the experiment into pounds at the following rate:

40 Tokens = 1 Pound

In this experiment, you will participate in 17 independent and identical (of the same form) auction rounds. In each round you will act as a bidder in an auction and compete for a single hypothetical object with other two participants in your group. Note that the first two rounds are practice rounds in which your earnings will not be counted for actual payoffs. The remaining 15 auction rounds are real and any positive or negative earnings will be counted for actual payoffs. If your balance during the experiment goes below zero, you will become inactive and be excluded for any remaining auction rounds and will receive only £5 participation fee at the end of the experiment.

An auction round

Next, we will describe in detail the process that will be repeated in all 17 rounds. Each round starts by having the computer randomly form three-participant groups. In each group, two participants are played by the computer (called computer participants), while the other participant is played by you.

In the beginning of each round, each participant will be assigned a *signal* that will be randomly drawn from the set of integer numbers of tokens between 0 and 100 (numbers not including decimals). That is, any number from the set {0, 1, 2, …, 100} will be equally likely to be drawn. A signal you will be assigned in each round is independent of signals other participants will be assigned and is independent of a signal assigned to you in any of the other rounds. This will be done by the computer.

You will only know your own signal but not signals of the other two computer participants. On the other hand, the computer participants will know not only their own signals but also the signals of other two participants.

The value of the object for each participant is determined by signals received by that participant and the other participants in the same group. Specifically, each participant's

Instructions English $k = 0$

Instructions

This is an experiment in the economics of decision-making. Research foundations have provided funds for conducting this research. Your earnings will depend partly on your decisions and partly on the decisions of the other participants in the experiment. If you follow the instructions and make careful decisions, you may earn a considerable amount of money.

At this point, check the name of the computer you are using as it appears on the top of the monitor. At the end of the experiment, you should use your computer name to claim your payments. At this time, you will receive £5 as a participation fee simply for showing up on time. In addition, you will receive £10 as an initial balance you will use in this experiment. Any positive or negative earnings incurred during the experiment will be added into this balance. Details of how you will make decisions will be provided below. During the experiment we will speak in terms of experimental tokens instead of pounds. Your payments will be calculated in terms of tokens and then exchanged at the end of the experiment into pounds at the following rate:

40 Tokens = 1 Pound

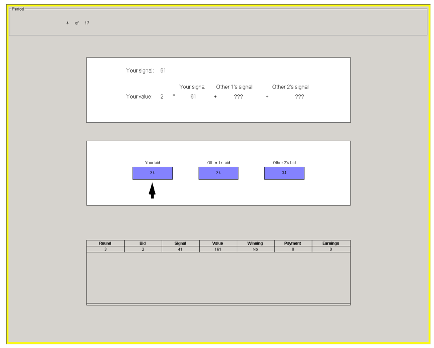
In this experiment, you will participate in 17 independent and identical (of the same form) auction rounds. In each round you will act as a bidder in an auction and compete for a single hypothetical object with other two participants in your group. Note that the first two rounds are practice rounds in which your earnings will not be counted for actual payoffs. The remaining 15 auction rounds are real and any positive or negative earnings will be counted for actual payoffs. If your balance during the experiment goes below zero, you will become inactive and be excluded for any remaining auction rounds and will receive only £5 participation fee at the end of the experiment.

An auction round

Next, we will describe in detail the process that will be repeated in all 17 rounds. Each round starts by having the computer randomly form three-person groups. The groups formed in each round depend solely upon chance and are independent of the groups formed in any of the other rounds. That is, in any group each active participant is equally likely to be chosen for that group. In a case where any other participant was excluded due to its negative balance, there is a chance that you may become inactive in a particular round when you are matched with that participant who was excluded.

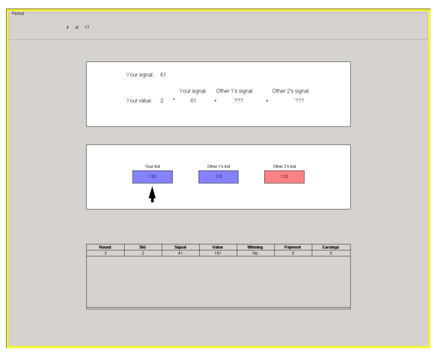
In the beginning of each round, each participant will be assigned a *signal* that will be randomly drawn from the set of integer numbers of tokens between 0 and 100 (numbers not including decimals). That is, any number from the set {0, 1, 2, …, 100} will be equally likely to be drawn. A signal you will be assigned in each round is independent of signals other participants will be assigned and is independent of a signal assigned to you in any of the other rounds. This will be done by the computer.

Attachment 1



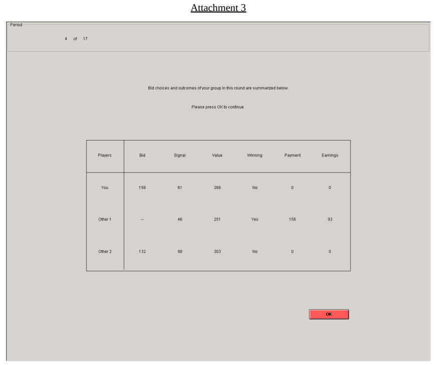
1

Attachment 2



2

Attachment 3



3

Figure 4.8: Attachments Instructions English $k = 0$

Instructions English $k = 1$

Instructions

This is an experiment in the economics of decision-making. Research foundations have provided funds for conducting this research. Your earnings will depend partly on your decisions and partly on the decisions of the other participants in the experiment. If you follow the instructions and make careful decisions, you may earn a considerable amount of money.

At this point, check the name of the computer you are using as it appears on the top of the monitor. At the end of the experiment, you should use your computer name to claim your payments. At this time, you will receive £5 as a participation fee simply for showing up on time. In addition, you will receive £10 as an initial balance you will use in this experiment. Any positive or negative earnings incurred during the experiment will be added into this balance. Details of how you will make decisions will be provided below. During the experiment we will speak in terms of experimental tokens instead of pounds. Your payments will be calculated in terms of tokens and then exchanged at the end of the experiment into pounds at the following rate:

40 Tokens = 1 Pound

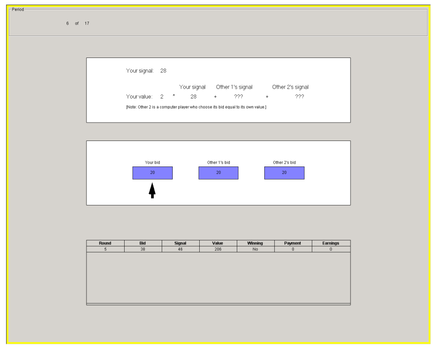
In this experiment, you will participate in 17 independent and identical (of the same form) auction rounds. In each round you will act as a bidder in an auction and compete for a single hypothetical object with other two participants in your group. Note that the first two rounds are practice rounds in which your earnings will not be counted for actual payoffs. The remaining 15 auction rounds are real and any positive or negative earnings will be counted for actual payoffs. If your balance during the experiment goes below zero, you will become inactive and be excluded for any remaining auction rounds and will receive only £5 participation fee at the end of the experiment.

An auction round

Next, we will describe in detail the process that will be repeated in all 17 rounds. Each round starts by having the computer randomly form three-participant groups. In each group, one participant is played by the computer (called a computer participant), while the other two participants are played by persons. The groups formed in each round depend solely upon chance and are independent of the groups formed in any of the other rounds. That is, in any group each active person is equally likely to be chosen for that group. In a case where any other person was excluded due to its negative balance, there is a chance that you may become inactive in a particular round when you are matched with that person who was excluded.

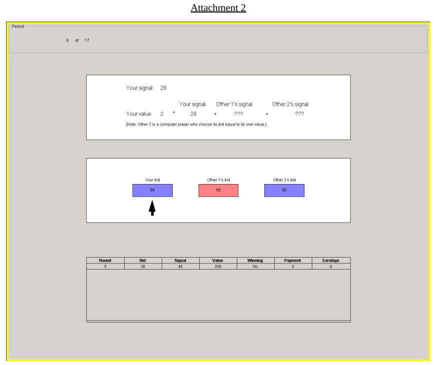
In the beginning of each round, each participant will be assigned a *signal* that will be randomly drawn from the set of integer numbers of tokens between 0 and 100 (numbers not including decimals). That is, any number from the set {0, 1, 2, …, 100} will be equally likely to be drawn. A signal you will be assigned in each round is independent of signals other participants will be assigned and is independent of a signal assigned to you in any of the other rounds. This will be done by the computer.

Attachment 1



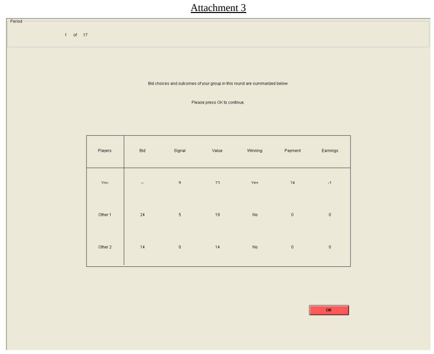
1

Attachment 2



2

Attachment 3



3

Figure 4.9: Attachments Instructions English any $k \in \{1, 2\}$

Instructions English $k = 2$

Instructions

This is an experiment in the economics of decision-making. Research foundations have provided funds for conducting this research. Your earnings will depend partly on your decisions and partly on the decisions of the other participants in the experiment. If you follow the instructions and make careful decisions, you may earn a considerable amount of money.

At this point, check the name of the computer you are using as it appears on the top of the monitor. At the end of the experiment, you should use your computer name to claim your payments. At this time, you will receive £5 as a participation fee simply for showing up on time. In addition, you will receive £10 as an initial balance you will use in this experiment. Any positive or negative earnings incurred during the experiment will be added into this balance. Details of how you will make decisions will be provided below. During the experiment we will speak in terms of experimental tokens instead of pounds. Your payments will be calculated in terms of tokens and then exchanged at the end of the experiment into pounds at the following rate:

40 Tokens = 1 Pound

In this experiment, you will participate in 17 independent and identical (of the same form) auction rounds. In each round you will act as a bidder in an auction and compete for a single hypothetical object with other two participants in your group. Note that the first two rounds are practice rounds in which your earnings will not be counted for actual payoffs. The remaining 15 auction rounds are real and any positive or negative earnings will be counted for actual payoffs. If your balance during the experiment goes below zero, you will become inactive and be excluded for any remaining auction rounds and will receive only £5 participation fee at the end of the experiment.

An auction round

Next, we will describe in detail the process that will be repeated in all 17 rounds. Each round starts by having the computer randomly form three-participant groups. In each group, two participants are played by the computer (called computer participants), while the other participant is played by you.

In the beginning of each round, each participant will be assigned a *signal* that will be randomly drawn from the set of integer numbers of tokens between 0 and 100 (numbers not including decimals). That is, any number from the set {0, 1, 2, …, 100} will be equally likely to be drawn. A signal you will be assigned in each round is independent of signals other participants will be assigned and is independent of a signal assigned to you in any of the other rounds. This will be done by the computer.

You will only know your own signal but not signals of the other two computer participants. On the other hand, the computer participants will know not only their own signals but also the signals of other two participants.

The value of the object for each participant is determined by signals received by that participant and the other participants in the same group. Specifically, each participant's

1

Chapter 5

Conclusion and Future Work

Regarding chapter 2 we provide an extension to discrete choice models with social interactions when there are asymmetries on the degree of peer influences and group unobservables. We find conditions under which a multinomial choice model presents unique equilibrium and show that, even in the presence of correlated effects at the group level, the endogenous and contextual effects are separately identified provided sufficient variation in the weighting matrix. In our application, our specific data allows us to investigate a particular historical period but our results might be relevant for social network effects in contemporary studies. Using contemporary datasets together with local networks based on geography not be the best measure of the relevant social group. This may lead to complicated patterns of interdependences in errors across individuals as well as make it difficult to assess counterfactuals (Manski 2000).

Relying on our historical period and the unique two-tier administrative system enables us to deal with self–selection into social group. We exploit the fact that public goods were provided at a higher tier and consequently affected location decision while community identity were still largely determined at a more local level. We find that social interactions are important in explaining industrial occupational choice and unemployment. Failing to account for group unobservables overestimates these effects.

Studying social interactions may help understand Inter-generational occupational (Borjas 1994, Munshi & Wilson 2008). This is a interesting question for future work.

On chapter 3 we provide an artefactual field experiment evidence on reference–dependent preferences and investment decisions on a sample of vulnerable small entrepreneurs from the developing world. In particular we find that having unfulfilled recent beliefs about a monetary outcome triggers risk–loving behaviour only on individuals who fell behind their more optimistic expectation, while promoting risk–averse behaviours on individuals which surpass their most pessimistic one.

Several authors have pointed out that expectations on future events can determine reference points. Our experimental evidence suggests that past expectations could affect also reference points determination in close relatedness to multiple reference points suggested by Köszegi & Rabin (2007), Sprenger (2010). We provide a model of stochastic reference points that replicates qualitatively most of the experimental results. Further experimental work is needed to fully understand the interactions between status quo, lagged beliefs and forward looking beliefs in determining reference points formation.

Lastly, the interdependent value auction model in chapter 4 explicitly studies the effects of insider information in standard auction formats. The structure of interdependent value enables us to investigate the implications on efficiency and revenue. We predict that the English auction has an efficient equilibrium, while the second-price auction is unable to avoid inefficiency. It is also shown that turning an outsider to an insider has a positive impact on the seller's revenues (reminiscent of Milgrom & Weber (1982*a*)'s linkage principle).

The experimental evidence supports the theoretical predictions on efficiency and revenues. Efficient allocation is more frequent in the English auction than in the second-price auction when insider information is present. Once realized signals are accounted for, average revenues of the English auction increase in the number of insiders. Despite the compliance of experimental data to theory, there is substantial evidence of naive bidding in line with other literature (Kagel & Levin 2011, 1995). Intriguingly, we find that the degree of naive bidding declines in the number of insiders in the English auction. This is something to ponder over in developing a behavioural explanation of subjects' bidding behaviour where imperfectly informed bidders become more wary the more insiders are present.

Bibliography

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

Aguirregabiria, V. & Mira, P. (2007), 'Sequential estimation of dynamic discrete games', *Econometrica* **75**(1), 1–53.

Akerlof, G. A. (1980), 'A theory of social custom, of which unemployment may be one consequence', *The quarterly journal of economics* pp. 749–775.

Akerlof, G. A. & Kranton, R. E. (2000), 'Economics and identity', *Quarterly journal of Economics* pp. 715–753.

Anderberg, D. & Andersson, F. (2007), 'Stratification, social networks in the labour market, and intergenerational mobility*', *The Economic Journal* **117**(520), 782–812.

Andersen, E. B. (1970), 'Asymptotic properties of conditional maximum-likelihood estimators', *Journal of the Royal Statistical Society. Series B (Methodological)* pp. 283–301.

Anderson, G. M. (1988), 'Mr. smith and the preachers: the economics of religion in the wealth of nations', *The Journal of Political Economy* pp. 1066–1088.

Anderson, S. P., De Palma, A. & Thisse, J. F. (1992), *Discrete choice theory of product differentiation*, MIT press.

Arellano, M. & Bonhomme, S. (2011), 'Nonlinear panel data analysis', *Economics* **3**.

Ashenfelter, O. & Graddy, K. (2003), 'Auctions and the price of art', *Journal of Economic Literature* **41**(3), 763–787.

Bayer, P., Ross, S. & Topa, G. (2008), 'Place of work and place of residence: Informal hiring networks and labor market outcomes', *Journal of Political Economy* **116**(6), 1150–1196.

Bayer, P. & Timmins, C. (2007), 'Estimating equilibrium models of sorting across locations', *Economic Journal* **116**, 353–374.

Benabou, R. (1993), 'Workings of a city: location, education, and production', *The Quarterly Journal of Economics* **108**(3), 619–652.

Berkelaar, A., Kouwenberg, R. & Post, T. (2004), 'Optimal portfolio choice under loss aversion', *Review of Economics and Statistics* **86**(4), 973–987.

Besley, T., Coate, S. & Guinnane, T. W. (2004), *History matters: economic growth, technology and demographic test*, Stanford University Press, chapter Ch 10 Incentives, information, and welfare: England"s new Poor Law and the workhouse test, pp. 245–270.

Bikhchandani, S. & Riley, J. G. (1991), 'Equilibria in open common value auctions', *Journal of Economic Theory* **53**(1), 101–130.

Birtles, S. (1999), 'Common land, poor relief and enclosure: The use of manorial resources in fulfilling parish obligations 1601-1834', *Past and Present* pp. 74–106.

Bisin, A., Moro, A. & Topa, G. (2011), The empirical content of models with multiple equilibria in economies with social interactions, Technical report, National Bureau of Economic Research.

Blau, D. M. & Robins, P. K. (1990), 'Job search outcomes for the employed and unemployed', *Journal of Political Economy* pp. 637–655.

Blume, L. E., Brock, W. A., Durlauf, S. N. & Ioannides, Y. M. (2010), *Handbook of Social Economics*, Vol. 1B, Cambridge: MIT Press, chapter Ch 18 Identification of social interactions, pp. 853–964.

Boone, J., Chen, R., Goeree, J. K. & Polydoro, A. (2009), 'Risky procurement with an insider bidder', *Experimental economics* **12**(4), 417–436.

Booth, C. (1897), *Life and Labour of the People of London*, Manchester University Press.

Borjas, G. J. (1994), Immigration, ethnic identity, and assimilation: The intergenerational transmission of immigrant skills, *in* 'Economic aspects of international migration', Springer, pp. 139–154.

Boskin, M. J. (1974), 'A conditional logit model of occupational choice', *The Journal of Political Economy* pp. 389–398.

Boucher, V., Bramoullé, Y., Djebbari, H. & Fortin, B. (2014), 'Do peers affect student achievement? evidence from canada using group size variation', *Journal of Applied Econometrics* **29**(1), 91–109.

Bowman, D., Minehart, D. & Rabin, M. (1999), 'Loss aversion in a consumption–savings model', *Journal of Economic Behavior and Organization* **38**, 155 – 178.

Bramoullé, Y. (2013), 'Comment to social networks and the identification of peer effects', *Journal of Business & Economic Statistics* **31**(3), 264–266.

Bramoullé, Y., Djebbari, H. & Fortin, B. (2009), 'Identification of peer effects through social networks', *Journal of Econometrics* **150**(1), 41–55.

Brock, W. & Durlauf, S. (2001), 'Discrete choice with social interactions', *Review of Economic studies* **68**(2), 235–260.

Brock, W. & Durlauf, S. (2002), 'A multinomial choice model with neighborhood effects', *American Economic Review* **92**(2), 298–303.

Brock, W. & Durlauf, S. (2006), Multinomial choice with social interactions, *in* L. Blume & S. Durlauf, eds, 'The economy as an evolving complex system', Vol. 3, Oxford University Press.

Brock, W. & Durlauf, S. (2007), 'Identification of binary choice models with social interactions', *Journal of Econometrics* **140**(1), 52–75.

Calvó-Armengol, A. (2004), 'Job contact networks', *Journal of Economic Theory* **115**(1), 191–206.

Calvó-Armengol, A. & Jackson, M. (2007), 'Networks in labor markets: Wage and employment dynamics and inequality', *Journal of Economic Theory* **132**(1), 27–46.

Campbell, C. M. & Levin, D. (2000), 'Can the seller benefit from an insider in common-value auctions?', *Journal of Economic Theory* **91**(1), 106–120.

Chamberlain, G. (1980), 'Analysis of covariance with qualitative data', *Review of Economic Studies* **47**(1), 225–238.

Chung, K.-S. & Ely, J. C. (2000), Efficient and dominance solvable auctions with interdependent valuations, Technical report.

Cook, D., Kieschnick, R. & McCullough, B. (2008), 'Regression analysis of proportions in finance with self selection', *Journal of empirical finance* **15**(5), 860–867.

Dasgupta, P. & Maskin, E. (2000), 'Efficient auctions', *Quarterly Journal of Economics* pp. 341–388.

Davis, J. (1988), *Reforming London: the London government problem, 1855-1900*, Clarendon Press Oxford.

Dyos, H. J. (1953), *Workmen's Fares in South London, 1860-1914*, University College of Leicester.

Engelbrecht-Wiggans, R., Milgrom, P. R. & Weber, R. J. (1983), 'Competitive bidding and proprietary information', *Journal of Mathematical Economics* **11**(2), 161–169.

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

Firth, J. F. B. (1888), *Reform of London government and city guild*, London: S. Sonnenschein, Lowery, and Co., Paternoster Square.

Friedman, B. M. (2011), 'Economics: A moral inquiry with religious origins', *The American Economic Review* pp. 166–170.

Gabrielsen, A. (1978), 'Consistency and identifiability', *Journal of Econometrics* **8**(2), 261–263.

Gaviria, A. & Raphael, S. (2001), 'School-based peer effects and juvenile behavior', *Review of Economics and Statistics* **83**(2), 257–268.

Goeree, J. K. & Offerman, T. (2002), 'Efficiency in auctions with private and common values: An experimental study', *American Economic Review* pp. 625–643.

Goldsmith-Pinkham, P. & Imbens, G. W. (2013), 'Social networks and the identification of peer effects', *Journal of Business & Economic Statistics* **31**(3), 253–264.

Gomes, F. (2005), 'Portfolio choice and trading volume with loss-averse investors*', *The Journal of Business* **78**(2), 675–706.

Granovetter, M. S. (1973), 'The strength of weak ties', *American Journal of Sociology* pp. 1360–1380.

Green, D. R. (1982), 'Street trading in london: a case study of casual labour, 1830-60', *The structure of nineteenth-century cities, Londres, Croom Helm* pp. 129–151.

Green, D. R. (1988), 'Distance to work in victorian london: a case study of henry poole, bespoke tailors', *Business History* **30**(2), 179–194.

Green, D. R. (1991), 'The metropolitan economy: Continuity and change, 1800–1939', *London–A New Metropolitan Geograph. Edward Arnold, London* .

Hansen, R. G. & Lott, J. R. (1991), 'The winner's curse and public information in common value auctions: Comment', *The American Economic Review* pp. 347–361.

Hanushek, E. A., Kain, J. F., Markman, J. M. & Rivkin, S. G. (2003), 'Does peer ability affect student achievement?', *Journal of applied econometrics* **18**(5), 527–544.

Heckman, J. J. & Honore, B. E. (1990), 'The empirical content of the roy model', *Econometrica: Journal of the Econometric Society* pp. 1121–1149.

Helmers, C. & Patnam, M. (2014), 'Does the rotten child spoil his companion? spatial peer effects among children in rural india', *Quantitative Economics* **5**(1), 67–121.

Hendricks, K. & Porter, R. H. (1988), 'An empirical study of an auction with asymmetric information', *The American Economic Review* pp. 865–883.

Hendricks, K., Porter, R. H. & Wilson, C. A. (1994), 'Auctions for oil and gas leases with an informed bidder and a random reservation price', *Econometrica: Journal of the Econometric Society* pp. 1415–1444.

Hoggart, K. & Green, D. R. (1991), *London: a new metropolitan geography*, Edward Arnold.

Hong, H. & Shum, M. (2003), 'Econometric models of asymmetric ascending auctions', *Journal of Econometrics* **112**(2), 327–358.

Hotz, V. J. & Miller, R. A. (1993), 'Conditional choice probabilities and the estimation of dynamic models', *The Review of Economic Studies* **60**(3), 497–529.

Hoxby, C. (2000), Peer effects in the classroom: Learning from gender and race variation, Technical report, National Bureau of Economic Research.

Johnson, J. H. & Pooley, C. G. (1982), *The structure of nineteenth-century cities.*, St. Martins Press.

Kagel, J. H., Harstad, R. M. & Levin, D. (1987), 'Information impact and allocation rules in auctions with affiliated private values: A laboratory study', *Econometrica: Journal of the Econometric Society* pp. 1275–1304.

Kagel, J. H. & Levin, D. (1986), 'The winner's curse and public information in common value auctions', *The American economic review* pp. 894–920.

Kagel, J. H. & Levin, D. (1999), 'Common value auctions with insider information', *Econometrica* **67**(5), 1219–1238.

Kagel, J. H. & Levin, D. (2011), *in* 'Handbook of Experimental Economics', Vol. 2, Princeton: Princeton University Press, chapter Auctions: a survey of experimental research, 1995–2010, pp. –.

Kagel, J. & Levin, D. (1995), *in* 'Handbook of Experimental Economics', Vol. 1, Princeton: Princeton University Press, chapter Bidding in common-value auctions: A survey of experimental research, pp. 501–585.

Kahneman, D. & Tversky, A. (1979), 'Prospect theory: An analysis of decision under risk', *Econometrica* **47**(2), 263–292.

Kasahara, H. & Shimotsu, K. (2012), 'Sequential estimation of structural models with a fixed point constraint', *Econometrica* **80**(5), 2303–2319.

Kim, J. (2003), *Studies of Informational Assymetries in Standard Auctions*, University of Wisconsin–Madison.

Kim, J. (2008), 'The value of an informed bidder in common value auctions', *Journal of Economic Theory* **143**(1), 585–595.

Kim, J. & Che, Y.-K. (2004), 'Asymmetric information about rivals' types in standard auctions', *Games and Economic Behavior* **46**(2), 383–397.

Kim, Y. & Loury, G. (2013), 'Social externalities overlap and the poverty trap', *Unpublished Manuscript* .

Kirchkamp, O. & Moldovanu, B. (2004), 'An experimental analysis of auctions with interdependent valuations', *Games and Economic Behavior* **48**(1), 54–85.

Knack, S. & Keefer, P. (1997), 'Does social capital have an economic payoff? a cross-country investigation', *The Quarterly journal of economics* pp. 1251–1288.

Köszegi, B. & Rabin, M. (2006), 'A model of reference–dependent preferences', *Quaterly Journal of Economics* **121**(4), 1133–1165.

Köszegi, B. & Rabin, M. (2007), 'Reference–dependent risk attitudes', *American Economics Review* **97**(4), 1047–1073.

Kramarz, F. & Skans, O. N. (2007), 'With a little help from my... parents? family networks and youth labor market entry', *Unpublished paper, CREST* .

Krishna, V. (2003), 'Asymmetric english auctions', *Journal of Economic Theory* **112**(2), 261–288.

Lawton, R. (1959), 'The daily journey to work in england and wales', *Town Planning Review* **29**(4), 241.

Lee, L.-f. (2007), 'Identification and estimation of econometric models with group interactions, contextual factors and fixed effects', *Journal of Econometrics* **140**(2), 333–374.

Lee, L.-f., Li, J. & Lin, X. (2014), 'Binary choice models with social interactions under heterogeneous rational expectations', *Review of Economic and Statistics (forthcoming)* .

Lin, X. (2010), 'Identifying peer effects in student academic achievement by spatial autoregressive models with group unobservables', *Journal of Labor Economics* **28**(4), 825–860.

Lind, B. & Plott, C. R. (1991), 'The winner's curse: Experiments with buyers and with sellers', *The American economic review* pp. 335–346.

Loomes, G. & Sugden, R. (1986), 'Disappointment and dynamic consistency in choice under uncertainty', *The Review of Economic Studies* **53**(2), 271–282.

Manski, C. F. (1993), 'Identification of endogenous social effects: The reflection problem', *The Review of Economic Studies* **60**(3), 531–542.

Manski, C. F. (2000), Economic analysis of social interactions, Technical report, National bureau of economic research.

Maskin, E. (1992), Auctions and privatization, *in* H. Siebert, ed., 'Privatization', JCB Mohr Publisher, pp. 115–136.

Milgrom, P. R. & Weber, R. J. (1982*a*), 'A theory of auctions and competitive bidding', *Econometrica: Journal of the Econometric Society* pp. 1089–1122.

Milgrom, P. & Weber, R. J. (1982*b*), 'The value of information in a sealed-bid auction', *Journal of Mathematical Economics* **10**(1), 105–114.

Moffitt, R. (2001), *Social dynamics*, Cambridge: MIT Press, chapter Policy interventions, low-level equilibria, and social interactions, pp. 45–82.

Munshi, K. & Wilson, N. (2008), Identity, parochial institutions, and occupational choice: Linking the past to the present in the american midwest, Technical report, National Bureau of Economic Research.

Myerson, R. B. (1981), 'Optimal auction design', *Mathematics of operations research* **6**(1), 58–73.

Neyman, J. & Scott, E. L. (1948), 'Consistent estimates based on partially consistent observations', *Econometrica: Journal of the Econometric Society* pp. 1–32.

Ospina, R. & Ferrari, S. (2010), 'Inflated beta distributions', *Statistical Papers* **51**(1), 111–126.

Ospina, R. & Ferrari, S. (2011), 'A general class of zero-or-one inflated beta regression models', *Computational Statistics & Data Analysis* .

Papke, L. & Wooldridge, J. (1996), 'Econometric methods for fractional response variables with an application to 401(k) plan participation rates', *Journal of Applied Econometrics* **11**(6), 619–632.

Patacchini, E. & Zenou, Y. (2012), 'Ethnic networks and employment outcomes', *Regional Science and Urban Economics* **42**(6), 938–949.

Patel, K. & Vella, F. (2013), 'Immigrant networks and their implications for occupational choice and wages', *Review of Economics and Statistics* **95**(4), 1249–1277.

Pendrill, C. (1937), *Old Parish Life in London.[With plates.]*., Oxford University Press.

Perry, M. & Reny, P. J. (2002), 'An efficient auction', *Econometrica* **70**(3), 1199–1212.

Pesendorfer, M. & Schmidt-Dengler, P. (2008), 'Asymptotic least squares estimators for dynamic games', *The Review of Economic Studies* **75**(3), 901–928.

Pesendorfer, M. & Schmidt-Dengler, P. (2010), 'Sequential estimation of dynamic discrete games: A comment', *Econometrica* **78**(2), 833–842.

Pooley, C. G. & Turnbull, J. (1997), 'Changing home and workplace in victorian london: the life of henry jaques, shirtmaker', *Urban History* **24**(02), 148–178.

Post, T., Van den Assem, M., Baltussen, G. & Thaler, R. (2008), 'Deal or no deal? decision making under risk in a large-payoff game show', *The American economic review* **98**(1), 38–71.

Riley, J. G. & Samuelson, W. F. (1981), 'Optimal auctions', *The American Economic Review* pp. 381–392.

Sacerdote, B. (2001), 'Peer effects with random assignment: Results for dartmouth roommates.', *Quarterly Journal of Economics* **116**(2), 681–704.

Schmiechen, J. A. (1984), *Sweated industries and sweated labor: The London clothing trades, 1860-1914*, Croom Helm London.

Shleifer, A. & Vishny, R. W. (1987), Management buyouts as a response to market pressure, *in* 'Mergers and acquisitions', University Of Chicago Press, pp. 87–102.

Siegmann, A. (2002), 'Optimal saving rules for loss–averse agents under uncertainty', *Economics Letters* **77**, 27 – 34.

Smith, R. M. (1904), *The religious life of London*, Hodder and Stoughton.

Snell, K. D. (2009), *Parish and belonging: community, identity and welfare in England and Wales, 1700-1950*, Cambridge University Press.

Solon, G., Page, M. E. & Duncan, G. J. (2000), 'Correlations between neighboring children in their subsequent educational attainment', *Review of Economics and Statistics* **82**(3), 383–392.

Spiegler, R. (2012), 'Monopoly pricing when consumers are antagonized by unexpected price increases: A "cover version" of the heidhues-koszegi-rabin model', *Economic Theory* **51**(3), 695–711.

Sprenger, C. (2010), An endowment effect for risk: Experimental tests of stochastic reference points, Technical report, working paper.

Stewart, C. (1900), *The Housing Question in London 1855-1900*.

Tanaka, T., Camerer, C. & Nguyen, Q. (2010), 'Risk and time preferences: Linking experimental and household survey data from vietnam', *American Economic Review* **100**, 557–571.

Topa, G. (2001), 'Social interactions, local spillovers and unemployment', *The Review of Economic Studies* **68**(2), 261.

Topa, G. (2011), Chapter 22 - labor markets and referrals, Vol. 1 of *Handbook of Social Economics*, North-Holland, pp. 1193 – 1221.
URL: *http://www.sciencedirect.com/science/article/pii/B9780444537072000050*

Topa, G. & Zenou, Y. (2014), 'Neighborhood versus network effects', *Handbook of Regional and Urban Economics* **4**.

Varadhan, R. & Gilbert, P. (2009), 'Bb: An r package for solving a large system of nonlinear equations and for optimizing a high-dimensional nonlinear objective function', *Journal of Statistical Software* **32**(4), 1–26.

Warnes, A. (1972), 'Estimates of journey-to-work distances from census statistics', *Regional Studies* **6**(3), 315–326.

Webb, S. & Webb, B. (1929), *English poor law history*, Vol. 8, Longmans, Green and Company.

Wilson, R. (1998), 'Sequential equilibria of asymmetric ascending auctions: The case of log-normal distributions', *Economic Theory* **12**(2), 433–440.

Yogo, M. (2008), 'Asset prices under habit formation and reference-dependent preferences', *Journal of Business & Economic Statistics* **26**(2), 131–143.