Occupational choice and social interactions: A Study of Victorian London*

José-Alberto Guerra† Myra Mohnen‡

This version: December 25, 2014
Latest version: http://jguerraforero.wix.com/joseaguerra

Abstract

This paper presents a multinomial choice model with social interactions and asymmetric influence. Individuals form rational expectations about the behaviour of peers by taking into account their characteristics and the strength of their ties. We show that, thanks to the asymmetries in peer influences, the effect from group members' expected behaviour (*endogenous effect*) and the effect from peers' characteristics (*contextual effect*) can be separately identified even when unobservables hit the group as a whole (*correlated effect*). We 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, social groups are delimited by ecclesiastical parish boundaries. Using census data that pins residential locations down to the street level, we 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 allow us to mitigate the self-selection bias. Our results show that social networks were important in determining occupational choice. Failing to allow for group unobservables 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.

*We are grateful to Syngjoo Choi, Aureo de Paula, Uta Schöenberg, Patrick Bayer, Lawrence Blume, Yann Bramoullé, Antonio Cabrales, Rachel Kranton, Imran Rasul, Martin Weidner, Yves Zenou and audiences at UCL, IFS, EDePo and Cambridge-INET Institute and HCEO Chicago Summer School in Social Economics for valuable comments. 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. Corresponding author: j.a.guerra.09@ucl.ac.uk (José-Alberto Guerra)

†Department of Economics, University College London. E-mail: j.a.guerra.09@ucl.ac.uk

‡Department of Economics, University College London. E-mail: myra.mohnen.09@ucl.ac.uk

"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

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, in the presence of correlated effects, the reflection 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, multiplicity of equilibria must be taken into account in order to get consistent and efficient estimates.[1]

Our aim in this paper is twofold: first we provide a discrete choice model extending Brock & Durlauf (2006) multinomial choice framework with group interactions to one with network interactions, and second we offer an empirical application on how interaction among members of a group affects their occupational decisions. At the core of our specification there is an interplay between an endogenous effect with respect to the local network and an exogenous one through public good provision (subsuming the correlated effect) 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. The network structure is introduced by a weighting matrix which captures closeness between each pair of the same group, thus imposing asymmetric influence between members. The provision of public goods, such as schooling, roads and sewerage maintainance, affects the cost of choosing an occupation by the characteristics of the area and its residents. We therefore allow for correlated effects at the group level which capture the effect of such ammenities but also 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

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

(PML) 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. 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, case in which the PML is known to behave poorly. Because the number of individuals per group is large, the incidental parameters problem is not severe as in short panel cases.[2] We confirm this in light of our simulations.

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, 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 groups. 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. This amounts to argue that when choosing locations, individuals care only about public ammenities. As parishes belonging to the same BW shared the same public goods provision, the Act made parish membership orthogonal to other unobservables that affect individual labour market decisions.

From our empirical investigations, we document spatial clustering of occupation in 1881 London even after controlling for neighbourhood effects, the latter may capture supply driven clustering of activities. 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.

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

Given the limitations of our data, we remain agnostic about potential mechanisms driving 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 paper 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.

As Bramoullé, Djebbari & Fortin (2009) we also explore the effect of social network structure on identification. As Lin (2010) we also explore various specifications of the spatial weights matrix and its effect on the estimates. 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, Kain, Markman & Rivkin 2003), rooms and dorms (Sacerdote 2001), neighbouring families (Bayer et al. 2008, Helmers & Patnam 2014, Solon, Page & Duncan 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 paper 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 paper 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 paper is also related to the literature on social capital (Knack & Keefer 1997).

The rest of the paper is organised as follows. Section 2 present the multinomial model with social interaction and asymmetric influence. Section 3.1 paints the historical background of London in the nineteenth century as the setting of our application while section 3.2 presents the the dataset. Section 4 presents our results and 5 investigates how robust they are. We finally summarise our findings and conclude in the last section.

2 Empirical Model

In this section, we present a model of occupational choice under incomplete information and network interactions. Then we prove 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 their notation.

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. $P_b = \{p \in P \mid b(p) = b\}$.[3] 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 $y \in \Omega = \{0, 1, \cdots, L-1\}$ expecting a market wage ω_y. Occupations are broadly defined so workers in different occupation are not perfect substitutes.

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 $i, j \in p$. That is, or 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 extent 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. Let us denote the distance between two individuals $i, j \in p$ as $d(i, j)$ we define a treshold δ as the maximum distance to which an individual can be influenced by other individual δ-away from him. Therefore, we define $w_{p,ij} = 1/|nei_{p,i}|$ if $d(i, j) \leq \delta$ and $w_{p,ij} = 0$ otherwise.

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. The network effect is embedded into $\phi_y(\omega_y, \mathbf{x}_i, \mathbf{s}^e_{py} \mid \mathbf{W}_p)$. $\mathbf{s}^e_{py}$ 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. The fact that individual characteristics enter the utility although one is controlling for expected wages may represent, as in Schmidt & Strauss (1975), preferences for or discrimination against certain occupations that depend on these observables which are not encompassed by prices.

[3] Where, with abuse of notation, $b(p)$ maps parish $p \in P$ to its corresponding administrative area $b \in B$.

[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.

$\phi(\cdot)$ 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) in that sense the peer effect partly stands for the option value of an occupation. It may also be related to the effect on utility of the availability for handling more/less efficiently occupation–specific problems when there are more workers of the same type close–by due to local complementarities/substitutions (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, $dim(\mathbf{z}_{p,i}) = S$. 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 $\mathbf{x}_j$. In that sense, the cost of following a given occupation is affected by public goods provided by the administrative area[5] or any effects explained mainly by peer's characteristics. For instance, an occupation–specific training school present in an administrative area may reduce the cost of following that type of occupation, but also a high incidence of neighbours from older cohorts may induce lower costs of following an occupation related to providing services to such population.

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

$$\max_{y} \phi_y \left(\omega_y, \mathbf{x}_i, \mathbf{s}^e_{py} \mid \mathbf{W}_p\right) - C_{b,y}\left(\mathbf{z}_{p_i}\right) + \nu_{p,y,i}, \tag{1}$$

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 $u_{p,y}$ which represents an information or demand type shock hitting the group as a whole. We assume both shocks are indepedent across occupations.

We parametrise $\phi_y(\omega_y, \mathbf{s}^e_{py} \mid \mathbf{W}_p) = k_y + \mathbf{x}_i\mathbf{c}_y + \mathbf{w}_{p,i}\mathbf{s}^e_{py}J_y$ and the term $-C_{b,y}\left(\mathbf{z}_{p,i}\right) = \mathbf{z}_{p,i}\mathbf{d}_y + \tau_{b,y}$. Notice we do not include wages ω_y directly in this approximation. Compared to the literature starting from Roy (1951)'s conceptual framework on occupational choice, our approach departs from the standard focus on foregone earnings as determinants of occupational choice as in Boskin (1974) or Heckman & Honore (1990). We are also shying away from any parental background affecting this choice (Borjas 1987). We do this in order to gain insight on the importance of social interaction channel and is mainly driven by the lack of such information on wages in our empirical application. 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 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.

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

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

[5] Such as schooling, roads improvements and sewerage coverage.

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

A.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 ($X_p \in \mathbb{I}_i$ 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 $\mathbb{I}_i$ 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

$$\mathbb{P}(y \in \arg\max_{y' \in \Omega} V(y'; p, \cdot) \mid \mathbb{I}_i) \equiv s_{py,i} = \frac{\exp\left(k_y + \mathbf{x}_i \mathbf{c}_y + \mathbf{z}_{p,i} \mathbf{d}_y + \mathbf{w}_{p,i} \mathbf{s}^e_{py} J_y + \tau_{b,y} + u_{p,y}\right)}{\sum_{y' \in \Omega} \exp\left(k_y + \mathbf{x}_i \mathbf{c}_{y'} + \mathbf{z}_{p,i} \mathbf{d}_{y'} + \mathbf{w}_{p,i} \mathbf{s}^e_{py'} J_{y'} + \tau_{b,y'} + u_{p,y'}\right)}. \tag{2}$$

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 $s^e_{py,j} = s_{py,j}$.

[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.

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

$$\mathbf{s}_{py} \equiv \begin{pmatrix} s_{py,1} \\ \vdots \\ s_{py,n_p} \end{pmatrix} = \begin{pmatrix} \frac{\exp\left(k_y+\mathbf{x}_1\mathbf{c}_y+\mathbf{z}_{p,i}\mathbf{d}_y+\mathbf{w}_{p,1}\mathbf{s}_{py}J_y+\tau_{b,y}+u_{p,y}\right)}{\sum\limits_{y'\in\Omega}\exp\left(k_{y'}+\mathbf{x}_1\mathbf{c}_{y'}+\mathbf{z}_{p,i}\mathbf{d}_{y'}+\mathbf{w}_{p,1}\mathbf{s}_{py'}J_{y'}+\tau_{b,y'}+u_{p,y'}\right)} \\ \vdots \\ \frac{\exp\left(k_y+\mathbf{x}_{n_p}c_y+\mathbf{z}_{p,i}\mathbf{d}_y+\mathbf{w}_{p,n_p}\mathbf{s}_{py}J_y+\tau_{b,y}+u_{p,y}\right)}{\sum\limits_{y'\in\Omega}\exp\left(k_{y'}+\mathbf{x}_{n_p}\mathbf{c}_{y'}+\mathbf{z}_{p,i}\mathbf{d}_{y'}+\mathbf{w}_{p,n_p}\mathbf{s}_{py'}J_{y'}+\tau_{b,y'}+u_{p,y'}\right)} \end{pmatrix} \tag{3}$$

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

$$\mathbf{S}_p = \Psi(\mathbf{S}_p, \boldsymbol{X}_p, \mathbf{Z}_p, \boldsymbol{W}_p; \boldsymbol{\theta}) \tag{4}$$

where $\boldsymbol{\theta} = \Big(k_y, c_y, d_y, J_y, (\tau_{b,y})_{b\in B}, (u_{p,y})_{p\in P}\Big)_{y\in\Omega}$

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

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

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

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 $\mathbf{m}_1 = \boldsymbol{\psi}_1(\hat{\mathbf{M}}, J, 0) = \frac{\exp(\mathbf{W}\mathbf{m}_1J)-1}{L-1+\exp(\mathbf{W}\mathbf{m}_1J)} \equiv \boldsymbol{\psi}_w(\mathbf{m}_1)$. 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, lies 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 $\frac{\partial \psi_w(\mathbf{m}_1')}{\partial m_{1,i}} = 0$ given $w_{ii} = 0$ and $\frac{\partial \psi_{w,i}(\mathbf{m}_1)}{\partial m_{1,j}} = \frac{\exp(\mathbf{w}_i\mathbf{m}_1 J) w_{ij} LJ}{L-1+\exp(\mathbf{w}_i\mathbf{m}_1 J)}$ for every $j \neq i$.

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

□

The previous result suggests 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. This is due to the assumption on the independence of errors which implies that with more alternatives, the choices are going to be largely influenced by the deterministic individual characteristics and there is less room for individuals coordinating or bunching together in some occupations due to $\mathbf{s}_{py}$

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 4 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)). It is worth highlithing that for the coming results on identification and estimation we do not require uniqueness.

2.2 Identification

There are two main threats to identification of the structural parameters $\boldsymbol{\theta}$ of the model 1–3. 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.

A.2. *Random assignment based on τ_b.*

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

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. *Under assumptions A.1, A.2, $L > 2$ 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 sufficient variation across rows so that, individuals within p interact only with a subset of individuals belonging to p.

then, for model described by 1-3, the true set of parameters $\boldsymbol{\theta} \setminus (J_y)_{y\in\Omega}$ are identified up to a normalisation while all $(J_y)_{y\in\Omega}$ are identified.

[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.

Proof. Given A.1, A.2 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. 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 (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 and sufficient within variation in at least one parish on its weighting matrix then the structural parameters are identified up to some normalization. In the linear–in–means case, Bramoullé et al. (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, whenever such difference in sizes across groups introduces non–linearities on the relation betwen peers actions and individual choices. In our case, given the non–linearities at the core of discrete choice modelling (see

[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.

Brock & Durlauf 2006), the requirement on the structure of the network is weaker and amounts to variation across rows.

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

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

$$L_N(\boldsymbol{Y} \mid \mathbf{X}, \boldsymbol{W}, \mathbf{S}; \boldsymbol{\theta}) = \frac{1}{N}\sum_{p\in P}\sum_{i\in p}\log\left[\frac{\sum_{y\in\Omega}\left(\exp\left(k_y + \mathbf{x}_i\mathbf{c}_y + \mathbf{w}_{p,i}\mathbf{X}_p\mathbf{d}_y + \mathbf{w}_{p,i}\mathbf{s}_{py}J_y + \tau_{b,y} + u_{p,y}\right)\mathbb{1}_{[y_i=y]}\right)}{\sum_{y'\in\Omega}\exp\left(k_y + \mathbf{x}_i\mathbf{c}_y + \mathbf{w}_{p,i}\mathbf{X}_p\mathbf{d}_y + \mathbf{w}_{p,i}\mathbf{s}_{py}J_y + \tau_{b,y} + u_{p,y}\right)}\right]. \tag{5}$$

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). For this method to solve the coherency problem due to multiple equilibria we need that, within a group, only one equilibrium is played in the data.

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, which clearly eliminates the need to solve for the fixed point problem while solving such pseudo likelihood maximization,[9] for any further step $t \geq 1$ such that the estimator at step t is

$$\hat{\boldsymbol{\theta}}^t = \arg\max_{\boldsymbol{\theta}\in\Theta} L_N(\boldsymbol{Y} \mid \mathbf{X}, \mathbf{W}, \hat{\mathbf{S}}^{t-1}; \boldsymbol{\theta}) \tag{6}$$

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

$$\hat{\mathbf{S}}^t = \boldsymbol{\Psi}(\hat{\mathbf{S}}^{t-1}, \mathbf{X}, \mathbf{W}; \hat{\boldsymbol{\theta}}^t) \tag{7}$$

and keep combining ML iteration with fixed–point updating until $\hat{\theta}^t$ is within a level of tolerance with respect to $\hat{\theta}^{t-1}$.[10]

[9] This two–step estimator using NR algorithm leads to asymptotically efficient estimates (Aguirregabiria & Mira 2007).

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

Hotz & Miller (1993), Pesendorfer & Schmidt-Dengler (2008) proposed alternative methods, while Pesendorfer & Schmidt-Dengler (2010) provide some global conditions on the fixed point mapping in which this iterative procedure fails to converge, with probability approaching 1, to the true parameters.

Due to this criticism, we also implement the Alternative Relaxation Method (i.e. *NPL-*Λ *algorithm* in Kasahara & Shimotsu (2012)) 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 (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) argue that[11] even when (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, by adding a set of dummies, 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 Andersen (1970) 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 (Chamberlain 1980). In section 4.1 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).

[11] See their Prop. 5

[12] See section 5.

[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 A.2 for further details.

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.

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 impoverished 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

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

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

Economic and social relevance of parishes

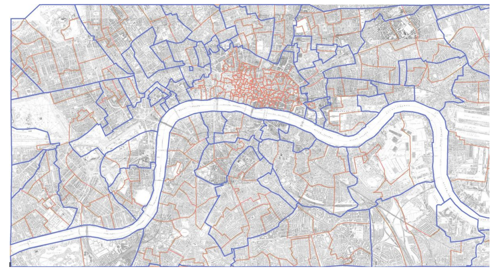
Despite losing importance in terms of civic responsibilities, Victorian age was a religious

[17] A detailed list of administrative areas in London can be found in the the appendix A.7.

[18] The Economist, June 19, 1869

[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.

(a) Ecclesiastical parishes (red)



(b) Board of Works and Vestries (blue)

Figure 1: Ecclesiastical and BW borders

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 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]

[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

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 social group on ecclesiastical parish boundaries provided, as we see below, 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. 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).

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.

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

dependant on weather conditions of the day.

Table 1: Distance to work

Period	Mean journey to work (km) 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)

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 A.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 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]

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 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 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.

4 Results

In the present section we include the results on the effects of social groups on occupational choice in Victorian London. We first study whether our estimation approach, using the PML

[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 A.1 we map these various characteristics. The south and eastern part of London 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 A.1, we provide comparison of these descriptives between the merged and the not merged datasets in order to assess the balance of our sample.

Table 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

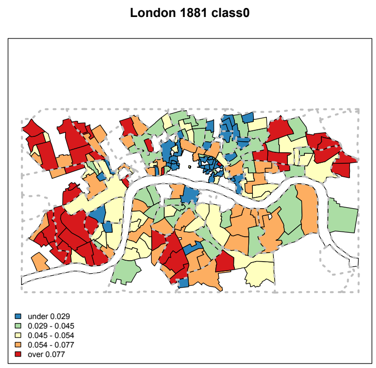
Notes: Std. dev in parenthesis. Sample includes only native working-age individuals (between 15 and 60) living in a parish which has a minimum of 30 residents within a BW which has at least two ecclesiastical parishes.

estimation with a fixed point subroutine is able to consistently estimate common parameters even in the presence of fixed effects by group.

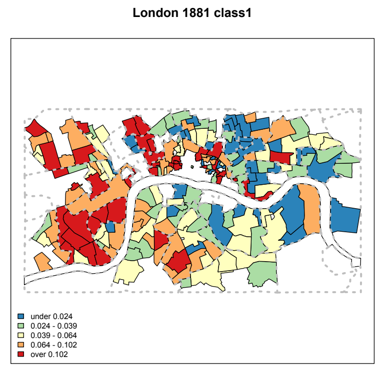
We then move to provide a group interactions/symmetric influence benchmark of the estimation and later our preferred specification with network interactions/asymmetric influence allowing for correlated effects at the group level.

4.1 Simulation to study the performance of the estimator

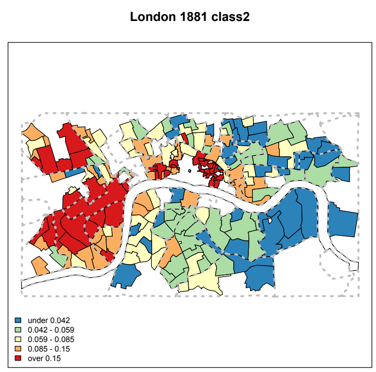
We take the empirical distribution of parish members and minimum number of neighbours living on the same street from our data set (see figure 10 in the Appendix) and randomly draw a duple $(n_p, \min_{i \in p} nei_i)$. We then simulate geographic points on $[0, U_{[1,2]}]^2$. We define $w_{p,i}$ as all 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 1,where we assume $K = S = 1$ and have to choose among five alternatives



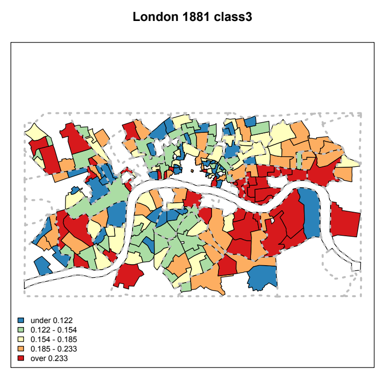
(a) Unempoyed



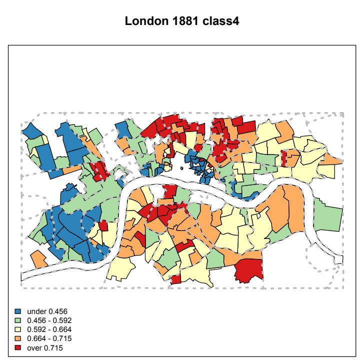
(b) Domestic Occupation



(c) Professional Occupation



(d) Commercial Occupation



(e) Industrial Occupation

Figure 2: Occupations per ecclesiastical parish.

$(L = 5)$.[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 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 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. $s_{p,y}$) 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).

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 6-7 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.

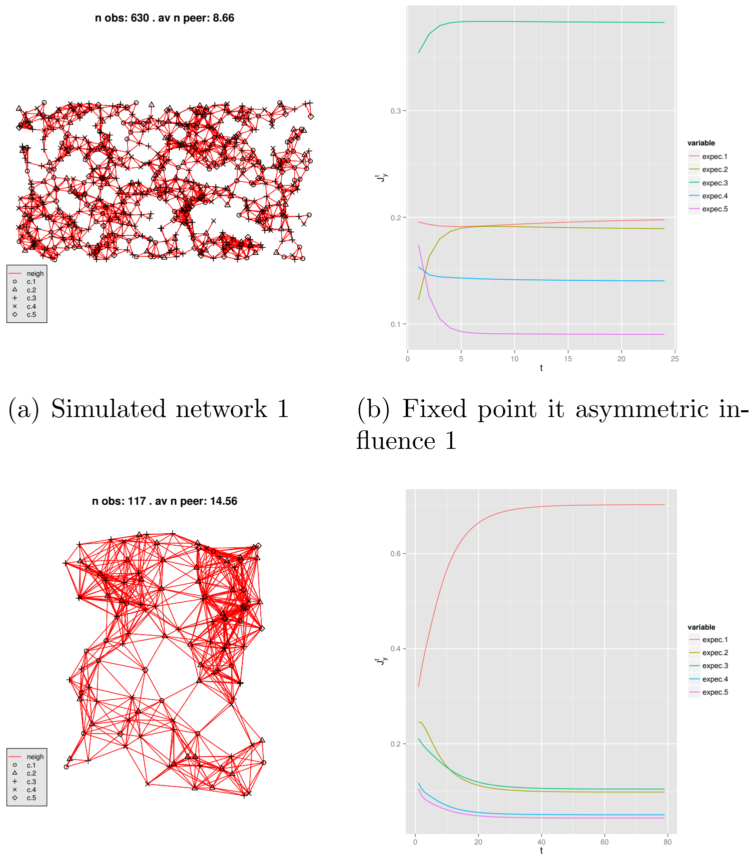
In Table 3 we show the endogenous parameter J_y estimates from the ML estimation assuming beliefs $s_{p,y}$ are not observed. We also explore the effects of measurement errors in the networks. In *Panel A*, the true network w_p is observed, while in *Panel B* the network is only partially observed w_p^* (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 es-

[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 12 and 13)

[27] In the Appendix A.3 we include corresponding estimates for the effect of an individual's own characteristics c (Table 11) and the contextual effect d (Table 10) effects as well as the estimates when both $s_{p,y}$ and w_p are observed (Table 9).



(a) Simulated network 1 (b) Fixed point it asymmetric influence 1

(c) Simulated network 2 (d) Fixed point it asymmetric influence 2

Figure 3: Some simulated network and beliefs fixed iteration per parishes with θ^0

timates 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%).

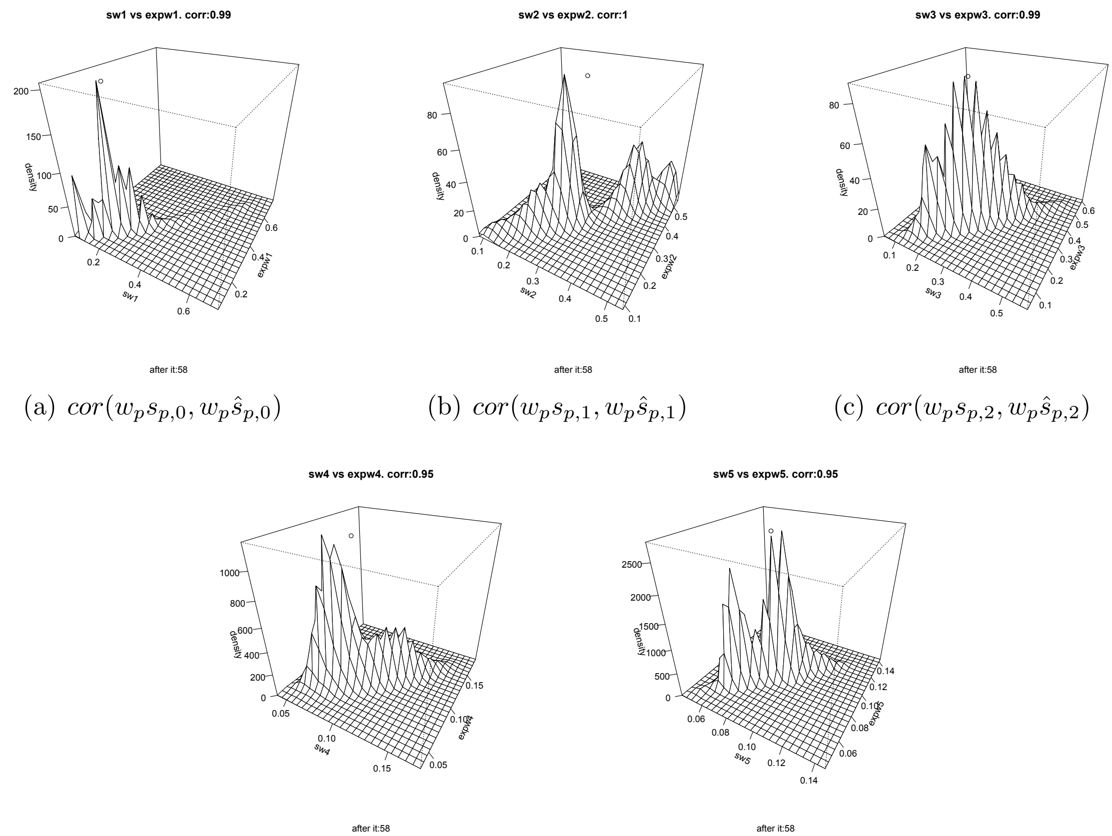
The results in *Panel B* are not as encouraging which should come as no surprise: as documented in the literature, whenever the true network is not observed, estimates are biased. A truncated network (in this case observing only up to observing 10 neighbours) generally lead to the underestimation of 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 13 in the appendix). This reinforces the importance of appropriately measuring the social group.

Table 3: Montecarlo simulations and estimation of J when $s_{p,y}$ is not observed

	(1) $J0 = 3.3$			(2) $J1 = 2.5$			(3) $J2 = 2$			(4) $J3 = 3.2$			(5) $J4 = 3.6$			$\min nei_i$	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*			
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 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.



(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 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

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 6-7 converge to the true parameters.[28]

4.2 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 (3) is now given by

$$s_{p,y} = \int \frac{\exp\left(k_y+x_i c_y+\bar{x}_p d_y+J_y s_{p,y}+\tau_{b,y}+u_{p,y}\right)}{\sum\limits_{y'\in\Omega}\exp\left(k_{y'}+x_i c_{y'}+\bar{x}_p d_{y'}+J_{y'} s_{p,y'}+\tau_{b,y}\right)} d\hat{F}_{\boldsymbol{x}|p}, \text{ for all } p \in P \tag{3a}$$

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).

We begin our econometric findings with the estimation of equation (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. Brock & Durlauf (2006) show that the structural parameters are identified provided there is no correlated effects, thus we assume no group unobservables. We also consider s_p^0 as being equal to the observed weighted average decision at group p ($\widehat{s}^0_{y,p} = \frac{1}{n_{p-1}}\sum_{j\in p}\mathbb{1}[y_j = y]$). 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

$$\theta^{naive} = \arg\max_{\theta} L(\boldsymbol{Y} \mid \mathbf{X}, \widehat{s}; \theta) \tag{8}$$

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 4 presents the endogenous effects J_y. Tables 14 in the appendix present the other estimates. The first three columns illustrate the naive estimates for J_y using equation (8)

[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).

while the last column uses the PML/FP using equation (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 4 that not incorporating the contextual effect is an important source of upward bias in our endogenous effect estimates. 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, $\frac{1}{n_{p-1}} \sum_{j \in p} s^t_{y,p}$) are smoother than the non–parametric consistent estimator s^0_p. 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 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 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.

[29] In Figure 14 panel *a)* we show the convergence of endogenous coefficients per iterations.

Table 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)
professional	9.522***	10.118***	8.247***	4.059***
	(0.279)	(0.357)	(0.418)	(0.955)
domestic	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 $s_{p,4}$.

4.3 Network interactions and asymmetric influence

In table 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 (4)) in the first two columns. We perform the PML/FP structural estimation described in expressions (6) and (7) in the last two columns. We successively add parish fixed effect to the basic specification.

We can now allow for group unobservables, an thus 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

[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 5.

share information. Failing to take the correlated effects into account can lead to serious upward bias.

Comparing column (4) of table 4 with column (3) of table 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 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.

[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 A.5

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 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
$w_p X_p$	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

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

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 1, we know that in the

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

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.

5 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.

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

In the first three columns of table 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 the network. For a distance of 0 *mts* results are very similar to our original same–street interactions.[33] 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.

[33] Notice in this particular case our parameters are identified provided that, within a parish, streets have different numbers of residents, thus guaranteeing that parish weighting matri's rows vary for two residents living in different streets.

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 6: Robustness checks estimation of endogenous effects J_y with asymmetric influence

vars	**PML/FP**						
	$w_p^{\delta=0}$	$w_p^{\delta=100}$	$w_p^{\|nei\|<10}$	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
	(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

5.2 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 (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

$$\Lambda^t = \left\{ \mathbf{\Psi}(\hat{\mathbf{S}}^{t-1}, \mathbf{X}, \mathbf{W}; \hat{\boldsymbol{\theta}}^t) \right\}^{\alpha} \hat{S}^{t-1^{1-\alpha}} \tag{9}$$

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

From the endogenous coefficients reported in Table 7 we notice that they do not change much compared to those reported in column (4) of Table 5 which is reassuring that our PML/FP á la Aguirregabiria & Mira (2007) estimates are consistent.

Table 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)
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 $s_{p,4}$.

5.3 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 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.[34] Therefore there was free mobility within a BW.

Figures 5 and 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[35] 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.

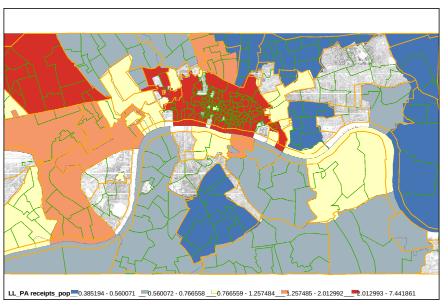


Figure 5: Total receipts (in £) per inhabitant by administrative area (BW/Vestries)

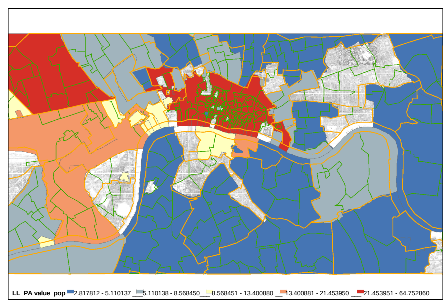


Figure 6: Assessable value (in £) per inhabitant by administrative area (BW/Vestries)

Source: The Economist Newspaper Ltd, London (1883)

[34]"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]

[35]see also figure 11 in appendix A.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.[36] 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 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 translates in this setting to

$$\lim_{r \to 0} Corr_\beta(\mathbb{E}_{\cdot|\beta_h}[Z_p], \mathbb{E}_{\cdot|\beta_r}[Z_p]) = \lim_{r \to 0} Corr_\beta(\mathbb{E}_{\cdot|\beta_h}[Z_p], \mathbb{E}_{\cdot|\beta_r}[Z_{p'}]) \tag{10}$$

If Z behaves as a random variable at the border β then condition (10) should hold. In contrast, if Z responds differently at either side of the border then condition (10) is no longer required to hold. In fact, if there is sorting patterns along parishes. exogenous

[36] "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).*

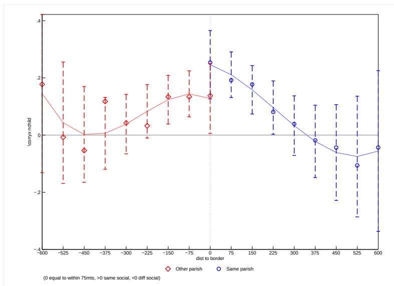
variables for individuals sharing the same social group should be more strongly correlated than for individuals belonging to different social groups. We would expect the correlation in the characteristics between individuals residing within the buffer zone and those living outside it, while still belonging 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 7 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.[37]

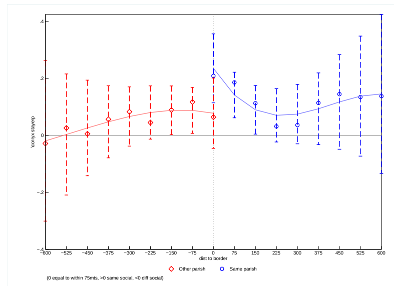
In figure 7 we see that none of the exogenous variables, apart from the share of married head of households, exhibit any discontinuities in their correlations while distinguishing by actual parish and neighbouring ones. Social interactions could be an important determinant for the marriage market, similarly 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 A.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.

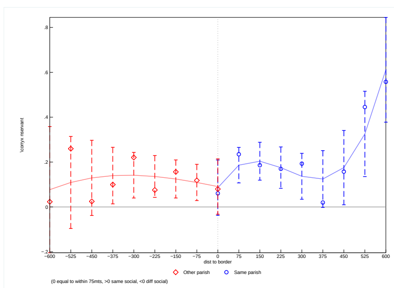
[37]We show the results for a buffer of $= 40$ mts and a degree third polynomial, but the results are robust to different buffers and polynomial degrees (see Appendix A).



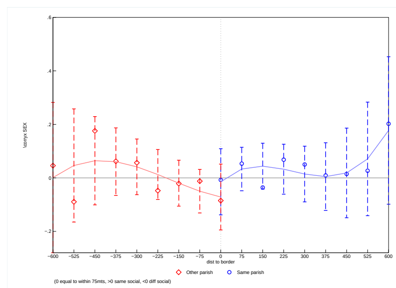
(a) number of children



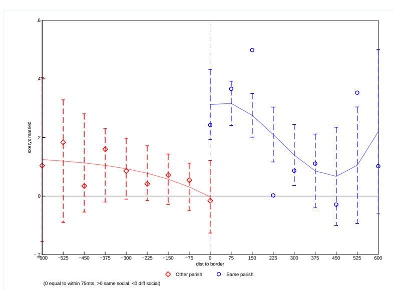
(b) parish birth=residence



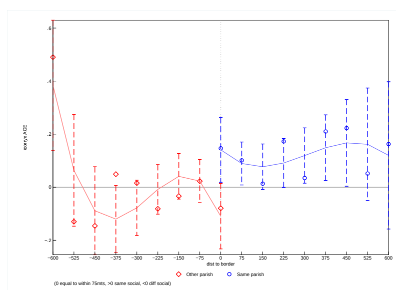
(c) number of servants



(d) male



(e) married



(f) age

Figure 7: Correlation in characteristics for $h = 40$ mts varying neighbours at 75 mts bins, polynomial degree 3

6 Conclusion

In this paper 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 paper 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 paper 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. Intergenerational 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 an interesting question for future work.

References

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**.

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.

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.

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.

Booth, C. (1897), *Life and Labour of the People of London*, Manchester University Press.

Borjas, G. J. (1987), 'Self-selection and the earnings of immigrants', *The American Economic Review* pp. 531–553.

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.

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.

Chamberlain, G. (1980), 'Analysis of covariance with qualitative data', *Review of Economic Studies* **47**(1), 225–238.

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.

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.

Goldsmith-Pinkham, P. & Imbens, G. W. (2013), 'Social networks and the identification of peer effects', *Journal of Business & Economic Statistics* **31**(3), 253–264.

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* .

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.

Hoggart, K. & Green, D. R. (1991), *London: a new metropolitan geography*, Edward Arnold.

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.

Kasahara, H. & Shimotsu, K. (2012), 'Sequential estimation of structural models with a fixed point constraint', *Econometrica* **80**(5), 2303–2319.

Kim, Y. & Loury, G. (2013), 'Social externalities overlap and the poverty trap', *Unpublished Manuscript* .

Knack, S. & Keefer, P. (1997), 'Does social capital have an economic payoff? a cross-country investigation', *The Quarterly journal of economics* pp. 1251–1288.

Kramarz, F. & Skans, O. N. (2007), 'With a little help from my... parents? family networks and youth labor market entry', *Unpublished paper, CREST* .

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 autoregessive models with group unobservables', *Journal of Labor Economics* **28**(4), 825–860.

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.

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.

Neyman, J. & Scott, E. L. (1948), 'Consistent estimates based on partially consistent observations', *Econometrica: Journal of the Econometric Society* pp. 1–32.

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.

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.

Roy, A. D. (1951), 'Some thoughts on the distribution of earnings', *Oxford economic papers* **3**(2), 135–146.

Sacerdote, B. (2001), 'Peer effects with random assignment: Results for dartmouth roommates.', *Quarterly Journal of Economics* **116**(2), 681–704.

Schmidt, P. & Strauss, R. P. (1975), 'The prediction of occupation using multiple logit models', *International Economic Review* pp. 471–486.

Schmiechen, J. A. (1984), *Sweated industries and sweated labor: The London clothing trades, 1860-1914*, Croom Helm London.

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.

Stewart, C. (1900), *The Housing Question in London 1855-1900*, London County Council by J. Truscott.

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, *in* A. B. Jess Benhabib & M. O. Jackson, eds, 'Handbook of Social Economics', 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**.

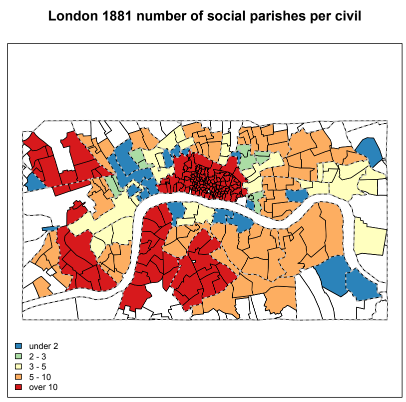
Varadhan, R. & Gilbert, P. (2009), 'Bb: An r package for solving a large system of nonlinear equations and for optimizing a high-dimensional nonlinear objective function', *Journal of Statistical Software* **32**(4), 1–26.

Warnes, A. (1972), 'Estimates of journey-to-work distances from census statistics', *Regional Studies* **6**(3), 315–326.

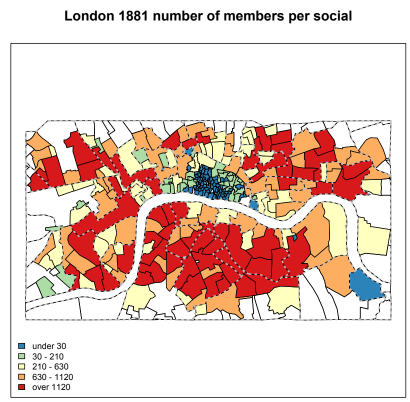
Webb, S. & Webb, B. (1929), *English poor law history*, Vol. 8, Longmans, Green and Company.

A Appendix

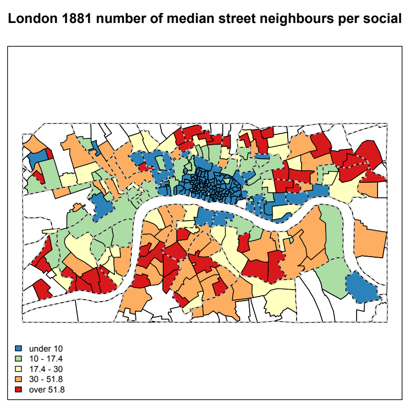
A.1 Further descriptives



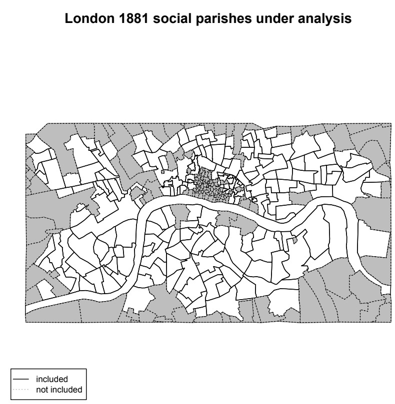
(a) $\mid P \mid$



(b) n_p

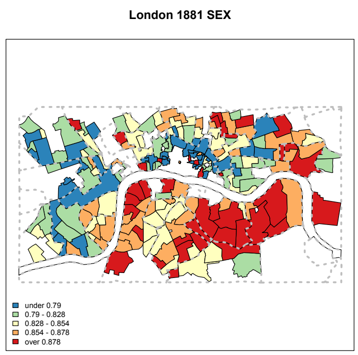


(c) nei_p

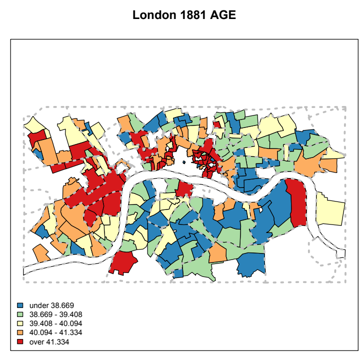


(d) Area under investigation

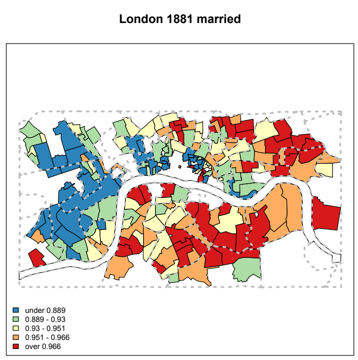
Figure 8: Social and civil parish: variation in $\mid P(b) \mid, n_p, nei(i)$



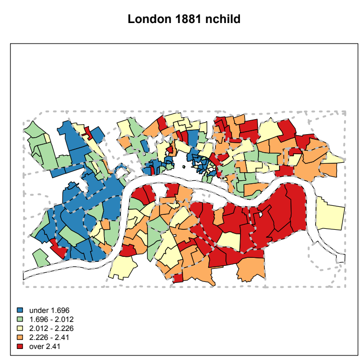
(a) Percentage of men per parish



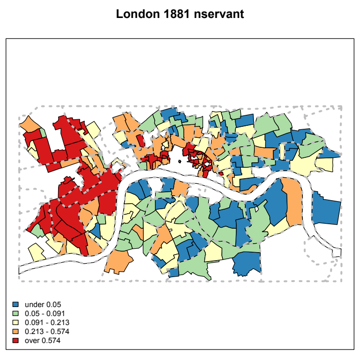
(b) Average age per parish



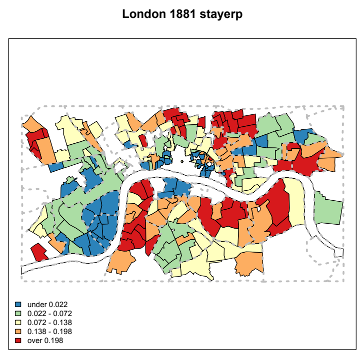
(c) Percentage of married individual per parish



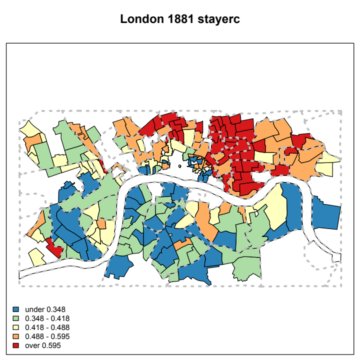
(d) Average number of children per parish



(e) Number of servants



(f) Percentage of residents in parish of birth



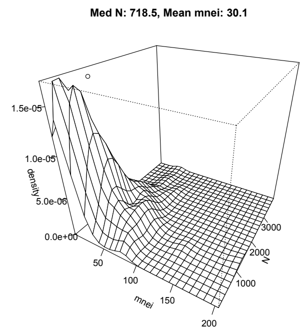
(g) Percentage of resident in county of birth

Figure 9: Descriptives by ecclesiastical parish

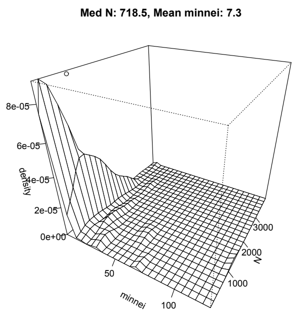
Table 8: Balanced Sample across merged observations in non institutional dwellings

	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

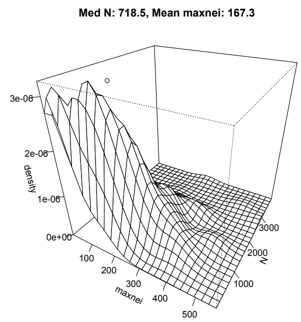
‡ *border parishes* are: Battersea, Bow, Bromley St Leonard, Brompton, Camberwell, St Dunstan Stepney/Mile End, Mile End New Town, Poplar, St George Hanover Square, St James Clerkenwell, St Leonard Shoreditch, St Luke Chelsea, St Luke Old Street, St Margaret Westminster, St Mary Abbots Kensington, St 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.



(a) median *nei*



(b) min *nei*



(c) max *nei*

Figure 10: Empirical distribution of parish members against statistics on street neighbours, London 1881

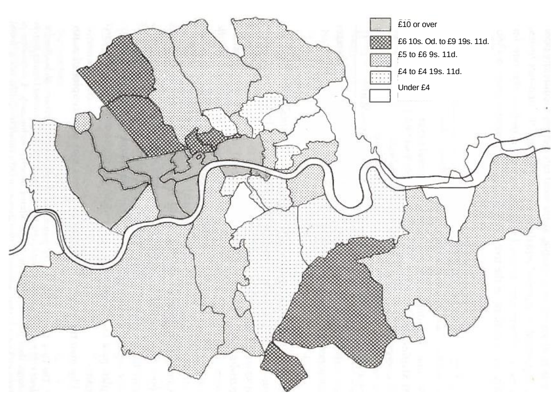


Figure 11: Local Areas: Rateable Value per Head, 1881

Source: Davis (1988)

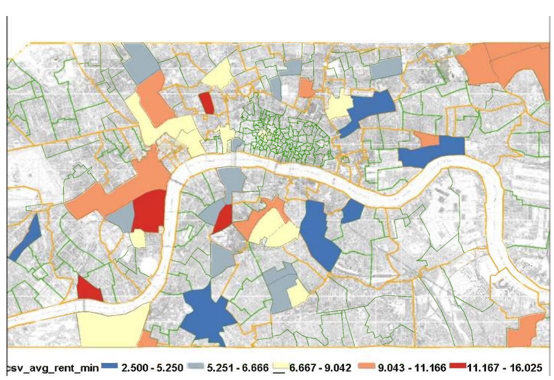


Figure 12: Avg. rent per room (in £, lower bound) by ecclesiastical parish

A.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'

$$\mathcal{L}_{N'} = \frac{1}{N'} \sum_{p \in P} \sum_{i \in p} \log \frac{\exp\left(\beta' \sum_{i,y} z_{p,iy} \mu_{p,iy}\right)}{\sum_{\lambda \in \Lambda_p} \exp\left(\beta' \sum_{i,y} z_{p,iy} \lambda_{iy}\right)}. \tag{11}$$

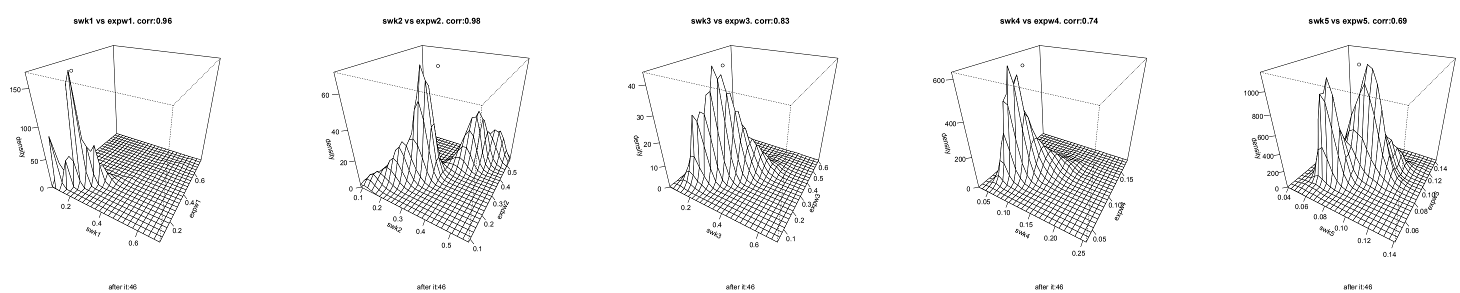
Where $\Lambda_p = \Big\{\lambda = (\lambda_{1,0}, \cdots, \lambda_{n_p,L-1}) \mid \lambda_{iy} = 0 \text{ or } 1, \sum_y \lambda_{iy} = 1, \sum_{i \in p} \lambda_{iy} = t_{py}, y = 1, \cdots,$
$L - 1\}$

A.3 Other Montecarlos results

Table 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 13: Correlation between $w_p s_{p,y}$ and $w_p^* \hat{s}_{p,y}$

Table 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 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 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 13: Montecarlo simulations and estimation of $J_y < 3.2$ for all $y \in \Omega$ and when $J_{y'} < 0$ 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) $J_y < 3.2$		$J0 = 2.8$			$J1 = 2$			$J2 = 1.5$			$J3 = 2.7$			$J4 = 3.1$			
$50^{\dagger}$	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*			
$100^{\ddagger}$	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) $J_y < 3.2$		$J0 = 2.8$			$J1 = 2$			$J2 = 1.5$			$J3 = 2.7$			$J4 = 3.1$			
$50^{\dagger}$	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*			
$100^{\ddagger}$	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
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.

A.4 Individual and contextual effects

Table 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)
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***

Continued on next page

vars	Homo, Naive			Homo, PML/FP	Hete, Naive		Hete, PML/FP	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	(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)
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***

Continued on next page

vars	Homo, Naive			Homo, PML/FP	Hete, Naive		Hete, PML/FP	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	(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)
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

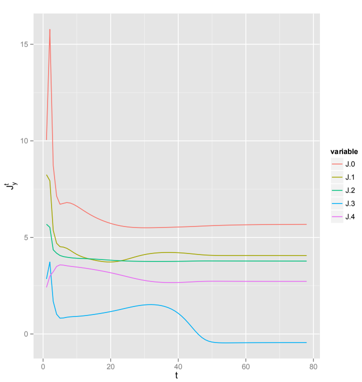
Continued on next page

vars	Homo, Naive			Homo, PML/FP	Hete, Naive		Hete, PML/FP	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	(0)	(0.723)	(0.854)	(1.075)	(0.301)	(0.325)	(0.313)	(0.322)
mean married_2		-0.301	-1.24	-1.849$^{\cdot}$	0.089	0.369	-0.479$^{\cdot}$	-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$^{\cdot}$	0.59$^{\cdot}$	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$^{\cdot}$	-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$^{\cdot}$	-0.146***	-0.123***	0.04	0.128*
	(0)	(0.087)	(0.1)	(0.131)	(0.034)	(0.036)	(0.05)	(0.052)
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

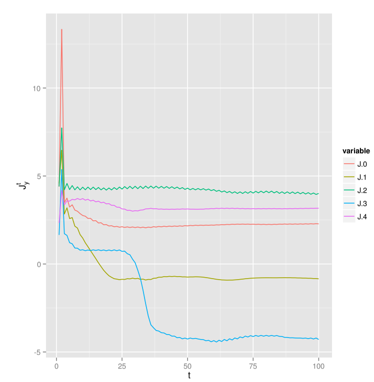
Continued on next page

vars	Homo, Naive			Homo, PML/FP	Hete, Naive		Hete, PML/FP	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	(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

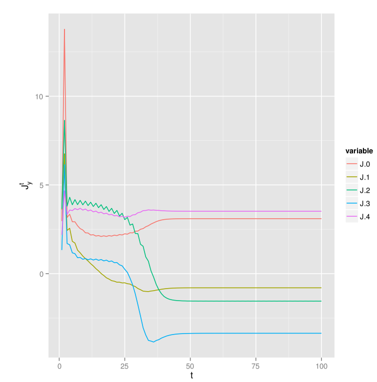
Notes: Standard errors in parentheses. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$



(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 14: Convergence of J_y^T for symmetric and asymmetric

A.5 Marginal effects

Starting from equations 2 and 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 but also an equilibrium effect through 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

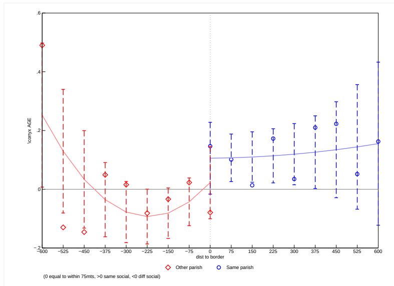
$$\begin{aligned}\tfrac{d\mathbb{P}_{y,i}}{dx_{k,i}} = \ & c_{k,y}\mathbb{P}_{y,i}\left(1-\textstyle\sum_{y'\in\Omega}\tfrac{c_{k,y'}}{c_{k,y}}\mathbb{P}_{y',i}\right) + \\ & J_y\mathbb{P}_{y,i}(1-\mathbb{P}_{y,i})\left[\textstyle\sum_{j\in nei_i} w_{ij}\left(w_{ji}d_{k,y}\mathbb{P}_{y,j}\left(1-\textstyle\sum_{y'\in\Omega}\tfrac{d_{k,y'}}{d_{k,y}}\mathbb{P}_{y'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

Table 15: Average Marginal Effects, endogenous and exogenous variables

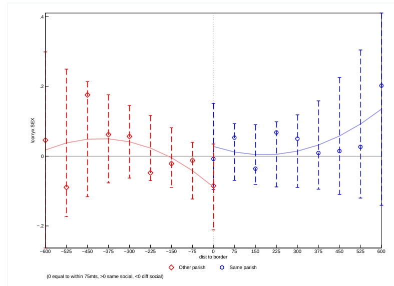
vars		**PML/FP**					
		unemployed	professional	domestic	commercial	industrial	
		(1)	(2)	(3)	(4)	(5)	(6)
Endogenous							
Estimated	$w_i \hat{s}_{y,i}$	5.37E-03	-2.41E-03	-1.83E-03	-2.46E-02	8.04E-02	
	sd $(w_i \hat{s}_{y,i})$	[0.039]	[0.048]	[0.049]	[0.054]	[0.106]	
Exogenous							sd var
	direct	-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.

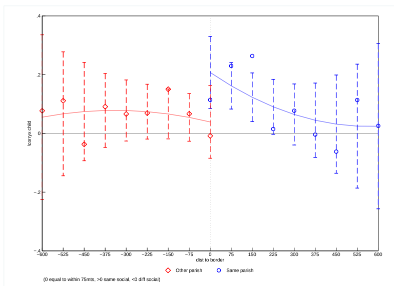
A.6 Discontinuities in correlations



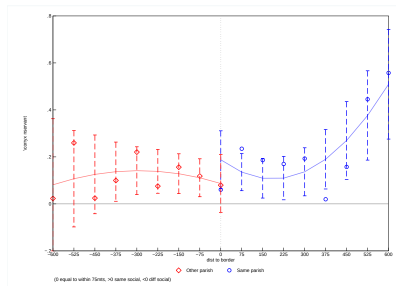
(a) age



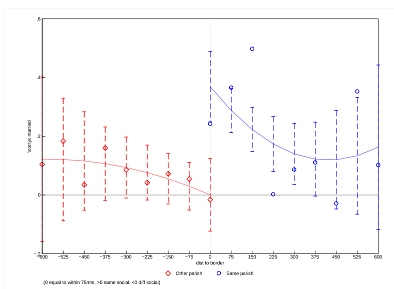
(b) male



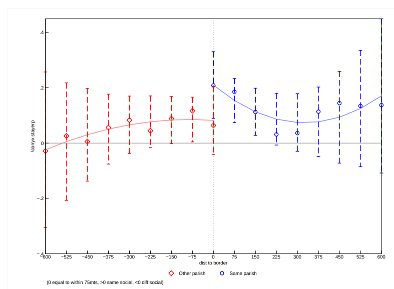
(c) children in HH



(d) number of servants

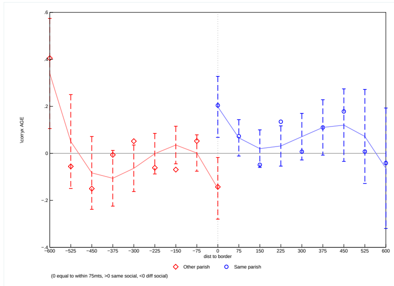


(e) married

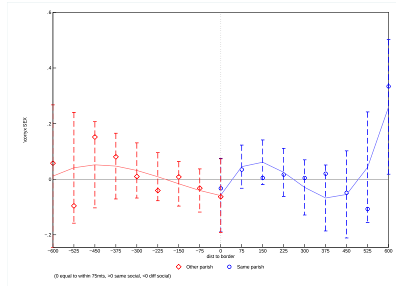


(f) parish birth=residence

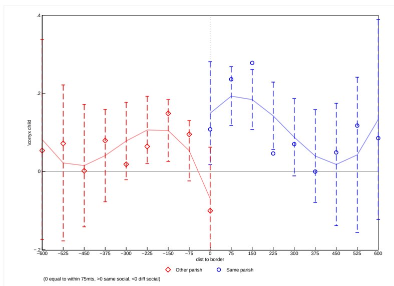
Figure 15: Correlation in endogenous variables for $h = 40$ mts varying neighbours at 75 mts bins, polynomial degree 2



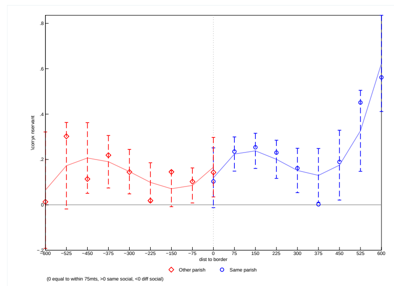
(a) age



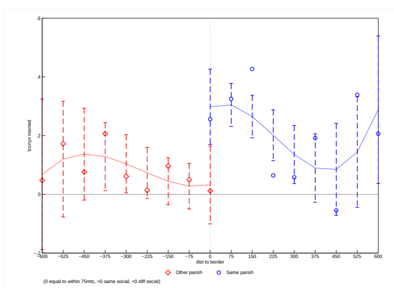
(b) male



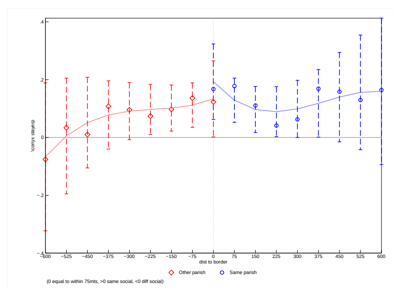
(c) children in HH



(d) number of servants



(e) married



(f) parish birth=residence

Figure 16: Correlation in exogenous variables for $h = 50$ mts varying neighbours at 75 mts bins, polynomial degree3

A.7 Administrative Areas

Table 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
	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

Continued on next page...

Present Borough	Metropolitan Borough	Board of Works	Civil Parishes	Ancient Parish
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
	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

Continued on next page...

Present Borough	Metropolitan Borough	Board of Works	Civil Parishes	Ancient Parish
		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
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`