Review of Economics and Statistics Just Accepted MS.
https://doi.org/10.1162/rest_a_00979

Multinomial Choice with Social Interactions: Occupations in Victorian London

José-Alberto Guerra and Myra Mohnen*

Abstract

We study the importance of social interactions on the occupational choice in Victorian London using a multinomial choice model within an incomplete social network. Individuals form heterogeneous rational expectations about their peers' behaviors taking into account their characteristics and the strength of their ties.

*Guerra: Department of Economics, Universidad de los Andes; Mohnen: Department of Economics, University of Essex and University of Ottawa, and Centre for Economic Performance, London School of Economics and Political Science. We thank the editor and four anonymous referees for their constructive comments. We are also grateful to Orazio Attanasio, Patrick Bayer, Oscar Becerra, Sandra Black, Lawrence Blume, Yann Bramoullé, Antonio Cabrales, Syngjoo Choi, Darwin Cortés, Marcela Eslava, Abhimanyu Gupta, Rachel Kranton, Aureo de Paula, Imran Rasul, Uta Schönberg, Andreas Uthemann, Martin Weidner, Yves Zenou and audiences at UCL, IFS, EDePo, Cambridge-INET Institute, ITAM, Universidad del Rosario, University of Melbourne, and Universidad de los Andes for helpful comments. We thank the Archaeology Department of the Museum of London for sharing their geographical referencing of London's historical map and Daniel Felipe Martinez Enriquez for assisting us in its extension.

Review of Economics and Statistics Just Accepted MS.
https://doi.org/10.1162/rest_a_00979

We show the conditions under which the endogenous, exogenous and correlated effects can be identified and a unique equilibrium can be established. Using a novel dataset, we proxy social groups by parish boundaries and strength of ties by geographic distances. Our results show the importance of the endogenous effects and reveal distinct effects by occupation.

JEL codes: C25, C31, J24, N93

1 Introduction

It is widely acknowledged that individuals are embedded in social networks that influence their behaviors and outcomes. Empirical analyses of social network effects have to face conceptual and data challenges. The structure and composition of an individual's social network are hard to measure. Analyses are further complicated by self-selection which can lead to a serious bias in the estimation of peer effect. In this paper, we study the importance of social interactions on the occupational choice of individuals in Victorian London. We address both empirical and theoretical issues using a novel dataset and a structural model approach.

We present a multinomial choice model with heterogeneous beliefs. In our model, individuals belong to a social group and interact with its members. When deciding their occupation, individuals form rational expectations about their peers' behavior taking into account their characteristics. Heterogeneity in these expectations are introduced through a weighting matrix which captures the strength of ties between peers. Correlated effects at the group level are included to capture potential shocks hitting the group as a whole. We show that the asymmetry in the influence of one's peers through the weighting matrix allows

Review of Economics and Statistics Just Accepted MS.
https://doi.org/10.1162/rest_a_00979

us to separately identify the endogenous effect (the influence of peers' behavior) and the exogenous effect (the influence of peers' attributes) from group unobservables. The conditions for unique equilibrium are also established. This framework may be applied to many areas involving local interactions and categorical outcomes such as criminal activities, modes of transport, or technology adoption. The structural parameters are estimated using a Recursive Pseudo Maximum Likelihood with an equilibrium fixed point subroutine following the Relaxation Method proposed by Kasahara and Shimotsu (2012) and Kasahara and Shimotsu (2018).

We then apply it to individuals' occupational choice in Victorian London. Combining a geo-referenced historical map of London with the full census of 1881, we construct a new dataset which allows us to locate individuals down to the street level. By exploiting the unique features of our historical setting and detailed dataset we overcome common empirical challenges. First, determining the appropriate reference group is difficult, especially in the modern world of high mobility and easy access to communication technologies. Researchers usually proxy the relevant group using some arbitrary metric of distance based on social and/or geographical proximity. Using a poor proxy for the true social group induces complicated patterns of interdependencies in errors across individuals. We present a novel definition of social group based on ecclesiastical parish boundaries. Parish membership was based on residency at a time when social networks were mostly local in nature and religion played a central role in community life (Booth, 1897). Consequently, ecclesiastical parish boundaries have the advantage of providing us with a more credible proxy for social groups. Second, computerization and the use of internet have been shown to play an important role in job search. Our historical setting suffers no such contamination in the role of informal

Review of Economics and Statistics Just Accepted MS.
https://doi.org/10.1162/rest_a_00979

contacts. Third, the fine spatial resolution of our data allows us to measure the strength of ties between parish members using geographical distances. Finally, the main threat to identification when examining peer effects is the sorting of individuals into social groups. We exploit the fact that London was operating under a two-tier administrative system. Parishes were grouped together to form local Board of Works (BW) or Vestries that were in charge of public good provision. By adding a fixed effect at the BW/Vestry level, we partly control for self-selection into groups.[1] We also include parish–level fixed effects to control for the possible confounding factors in the form of local industry shocks or a local priest who might encourage parishioners to share information.

Our results highlight the importance of social networks on occupational choice. There is a positive effect in the expected share of peers within a social group employed in a certain occupation on an individual's choice to follow the same occupation, regardless of the type of occupation. This reflects local complementarities stemming from learning spillovers, aspiration formation, or social capital. Indeed, peers can shape the beliefs of other parish members by sharing information about job opportunities or by sharing their experience about a par-

[1] In the absence of random peer groups, 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 individual choices cannot narrow their preferences down to the smaller preferred unit. In the case of class-school choice, families can decide which school to send their children to but cannot decide which class they will be part of. In our context, we argue that families decide where to live based on the amenities provided at the BW/Vestry level and cannot choose their exact locations due to the tight housing market. Nevertheless, it is important to acknowledge that our approach does not allow for non-random selection of individuals into groups.

Review of Economics and Statistics Just Accepted MS.
https://doi.org/10.1162/rest_a_00979

ticular occupation which affects aspiration. Alternatively, social capital can affect beliefs through social norms and/or peer pressure. Moreover, we uncover a large variation in the size of these endogenous effects depending on the type of occupation. A one standard deviation increase in the expected share of peers leads to an increase in the likelihood of choosing the same occupation ranging from 3.7% in the case of industrial artisan to a mere 0.44% in the case of professional occupation. This stark difference might be explained by the fact that the relevant peers for professionals operate at a higher level than the parish. More generally, we show that the endogenous effects are more precisely estimated under heterogeneous rational expectations than under the homogeneous case. Moreover, ignoring possible unobserved shocks at the group level might bias peer effect estimates.

Our results remain robust to alternative explanations, specifications, and estimator refinements. We provide evidence that our results are robust to the sorting of individuals into streets within parish. We also control for the influence of family members by creating a new dataset that tracks individuals from the 1851 to the 1881 census thus allowing us to identify the occupation of fathers and brothers. Further robustness checks include specifications looking at the age of individuals and alternative measures of strength of ties. Finally, a bias-correction estimation is implemented to account for fixed-effects in nonlinear models.

Assessing whether and to what extent social interactions influence occupational choice is important given the role of occupational structure in the process of development through the distribution of income and wealth (Banerjee and Newman, 1993). Our results yield a number of insights. Social networks can offer a possible explanation for the observed spatial clustering of occupation (Bayer et al., 2008) and the inequality patterns within a city (Glaeser et al., 2009). Moreover, the composition of social reference groups may produce a

Review of Economics and Statistics Just Accepted MS.
https://doi.org/10.1162/rest_a_00979

misallocation of talent and resources since contacts can help find job but not necessarily in the occupation where workers are most productive. Our findings therefore brings support to the view that social groups can be a determinant of cross country GDP and productivity differences (Murphy et al., 1991). Finally, our results provide a novel interpretation for the observed low intergenerational mobility and persistence of segregation by occupation (Solon, 1999). Intergenerational mobility may remain low because workers seek to use their inherited social connections to find jobs more easily (Borjas, 1994).

The model we propose builds on a growing number of theoretical models incorporating social interactions. In a seminal paper, Manski (1993) proposes a linear–in–means model with social interactions within a complete network. Individuals belong to a group, interact with all the other members within the group, and form rational expectations regarding their peers' behavior. Consequently, individuals are equally influenced by all members of their own group and form homogeneous rational expectations based on the group–level behavior. Blume et al. (2015) generalize this results to linear social interactions models where the actions have a continuous support. The homogeneous rational expectation framework is adopted by Brock and Durlauf (2001) in a binary choice model and Brock and Durlauf (2002, 2006) in a multinomial choice model. In the case of an incomplete network where individuals weigh their interactions by the strength of their ties, individuals may take into account their peers' characteristics thereby forming heterogeneous rational expectations. We provide such an extension and show that the conditions for a unique equilibrium are not more stringent than in the case of homogeneous expectations. Moreover, we find that, even in the presence of correlated effects, the endogenous parameters are identified for all the alternatives if there is sufficient variation in the network within group. Another closely related paper is

Review of Economics and Statistics Just Accepted MS.
https://doi.org/10.1162/rest_a_00979

Lee et al. (2014) which develops a heterogeneous rational expectations within a binary choice model. Not only do we focus on a multinomial choice model, but our empirical application and estimation strategy also differ from theirs.

Our paper also contributes to the empirical literature that focuses on the role of social networks for labor market outcomes. There is considerable evidence that social contacts are important for labor market outcomes (Ioannides and Loury, 2004). Most of the literature has assessed the role of social contacts on employment status, unemployment duration and wages. The part of this literature which is closely related to ours emphasizes network effects as neighborhood or ethnic group effects. The evidence suggests that networks defined by geographical proximity have a positive influence on employment. Bayer et al. (2008) use micro-level census data for Boston and find that residing in the same block raises the probability of sharing the work location by 33%, consistent with local referral effects. These neighborhood effects are particularly strong in the case of ethnic groups. Patacchini and Zenou (2012) focus on ethnic minorities in England and find that the higher the percentage of a given ethnic group living nearby, the higher the probability of finding a job through social contacts. Neighborhood effects however have for the most part ignored the micro-structure of connections due to data limitation. Thanks to our historical setting and unique dataset, we define social groups using religious and political borders and proxy for the strength of ties between individuals using geographical proximity.

The network literature on occupational choice is relatively scant. Munshi (2003) explores the influence of the occupational choice of nineteenth century immigrants on those of current immigrants from the same country through transmission of job information. He finds that community-based social interactions improve labor market outcomes among migrants. Patel

Review of Economics and Statistics Just Accepted MS.
https://doi.org/10.1162/rest_a_00979

and Vella (2013) find that new immigrants are more likely to choose the same occupation previous immigrants from the same country have chosen and enjoy a large and positive effect on their earnings. Most studies introducing spillovers from social networks into occupational choice take them as exogenous. For instance, in Cicala et al. (2017), exogenous spillovers change relative benefits from different activities. Albornoz et al. (2017) introduce endogenous spillovers where individuals choose whom to interact with and then choose the strength of their ties and productive efforts. To this body of work we add by systematically measuring the magnitude of endogenous spillovers by occupation category.

The rest of the paper is organized as follows. Section 2 presents the multinomial choice model and identification results. Section 3 shows the estimation strategy. We apply our model in section 4 to occupational choices in nineteenth century London. We present our results on the influence of peers on occupational choice and explore alternative explanations. We finally perform several robustness checks. In section 5 we conclude.

2 Empirical Model

Consider individual i embedded in a group g faced with a multinomial choice $y \in \mathcal{Y} = \{0, 1, \ldots, L-1\}$. There are G groups in the population, each $g \in G$ is characterized by a realization n_g of the random variable determining the size of the group. A row-normalized weighting matrix $\mathbf{W}_g$ (also known as an adjacency matrix), with entry $w_{g,ij} \forall i, j \in g$, represents the strength of the tie between i and j. We assume no self–influence (i.e. $w_{g,ii} = 0$). The $1 \times n_g$ vector $\mathbf{w}_{g,i}$ determines the weight individual i, belonging to group g, places on

Review of Economics and Statistics Just Accepted MS.
https://doi.org/10.1162/rest_a_00979

each member of her group.[2]

Individual i is characterized by a vector $\mathbf{x}_{g,i}$ ($dim(\mathbf{x}_{g,i}) = K$). Taking group membership as given, i chooses outcome $y \in \mathcal{Y}$ such as to maximize a pay–off function $V : \mathcal{Y} \to \mathbb{R}$, within a random utility maximization framework (McFadden, 1974). The individual utility consists of four components: a deterministic private utility $u(\cdot)$, a "social" utility $S(\cdot)$ that depends on group characteristics and outcomes, a group–level effect $\nu_{g,y}$, and a random private utility $\epsilon_{g,y,i}$ (Brock and Durlauf, 2006; Soetevent and Kooreman, 2007).

$$\max_{y \in \mathcal{Y}} u(y, \mathbf{x}_{g,i}) + S\left(y, \mathbf{X}_g, \mathbf{p}_{gy}^{e,i}, \mathbf{W}_g\right) + \nu_{g,y} + \epsilon_{g,y,i}, \tag{1}$$

where $\mathbf{X}_g$ is the $n_g \times K$ matrix with i^{th}-row equal to $\mathbf{x}_{g,i}$. $\nu_{g,y}$ and $\epsilon_{g,y,i}$ are unobservable to the econometrician. Let $\mathbb{I}_i = (\mathbf{W}_g, \mathbf{X}_g, \nu_{g,y} \forall y \in \mathcal{Y}, \epsilon_{g,y,i} \forall y \in \mathcal{Y})$ denote the amount of information available to individual $i \in g$ before choosing. Let $p_{gy,j}^{e,i}$ be the belief that individual $i \in g$ has about individual $j \in g$ taking action $y \in \mathcal{Y}$ and $\mathbf{p}_{gy}^{e,i} \equiv \left(p_{gy,j}^{e,i} \forall j \in g\right)$. Note that such a belief has superscript i because it potentially depends on $(\epsilon_{g,y,i} \forall y \in \mathcal{Y})$. We assume:

A0.a. *Incomplete information and rational expectations:* Individual $i \in g$ forms rational expectations regarding the behavior of members of her group. That is, $p_{gy,j}^{e,i}$ coincides with the objective belief for each $j \in g$, denoted by $p_{gy,j}^{i}$.

A0.b. *Linearity of private and social utilities:*

(i) $u(y, \mathbf{x}_{g,i}) = \alpha_y + \mathbf{x}_{g,i}\boldsymbol{\beta}_y$;

[2] With some abuse of notation, g is used as the social group index and the set of all individuals belonging to that group, and G is used to denote the total number of groups and the collection of groups.

Review of Economics and Statistics Just Accepted MS.
https://doi.org/10.1162/rest_a_00979

(ii) $S\left(y, \mathbf{X}_g, \mathbf{p}_{gy}^{e,i}, \mathbf{W}_g\right) = \mathbf{w}_{g,i}\mathbf{X}_g\boldsymbol{\delta}_y + \mathbf{w}_{g,i}\mathbf{p}_{gy}^{e,i}\gamma_y.$

A0.c. *Distribution of private utility shocks:*

(i) $\epsilon_{g,y,i}$ are independent and identically distributed across and within groups g with known Gumbel distribution function $F_\epsilon(\epsilon_{g,y,i} < \epsilon) = \exp(-\exp(-\epsilon))$. Hence, $p_{gy,j}^i$ can be denoted by $p_{gy,j}$.[3]

(ii) $(\epsilon_{g,y,i}\ \forall y \in \mathcal{Y},\ \forall i \in g) \perp (\mathbf{x}_{g,i}\ \forall i \in g, \mathbf{W}_g, \nu_{g,y} \forall y \in \mathcal{Y})$.

Notice assumption *A0.c.ii* requires that $(\epsilon_{g,y,i}\ \forall y \in \mathcal{Y},\ \forall i \in g) \perp \mathbf{W}_g$ which suggests that, once we account for group fixed effects, there are no unobservables among peers that explain whom they would like to interact with. Imposing assumptions $A0$, one can derive from expression (1) that agent i, belonging to social group g, chooses y with probability given by

$$p_{gy,i} = \frac{\exp(\alpha_y + \mathbf{x}_{g,i}\boldsymbol{\beta}_y + \mathbf{w}_{g,i}\mathbf{X}_g\boldsymbol{\delta}_y + \mathbf{w}_{g,i}\mathbf{p}_{gy}\gamma_y + \nu_{g,y})}{\sum_{y' \in \mathcal{Y}} \exp\left(\alpha_{y'} + \mathbf{x}_{g,i}\boldsymbol{\beta}_{y'} + \mathbf{w}_{g,i}\mathbf{X}_g\boldsymbol{\delta}_{y'} + \mathbf{w}_{g,i}\mathbf{p}_{gy'}\gamma_{y'} + \nu_{g,y'}\right)}. \quad (2)$$

In Manski (1993)'s terms, we wish to know whether we see correlations among peers because they share the same sources of information $\nu_{g,y}$ (correlated effect), because they share individual characteristics as a result of self-selection $\boldsymbol{\delta}_y$ (contextual or exogenous effect), or because they learn from one another's behavior γ_y (endogenous effect).

2.1 Equilibrium

Individuals maximize expected utility and each one has consistent beliefs about the choices of other parish members given by vector $\mathbf{p}_{gy}$. Therefore, the Bayes-Nash equilibrium to the individual choice vector $\mathbf{p}_{gy}$ is the fixed point solution to expression

[3] By mutual independence of taste shocks, $p_{gy,j}^i$ is independent of $(\epsilon_{g,y,i} \forall y \in \mathcal{Y})$.

Review of Economics and Statistics Just Accepted MS.
https://doi.org/10.1162/rest_a_00979

$$\mathbf{p}_{gy} \equiv \begin{pmatrix} p_{gy,1} \\ \vdots \\ p_{gy,n_g} \end{pmatrix} = \begin{pmatrix} \frac{\exp(\alpha_y+\mathbf{x}_{g,1}\boldsymbol{\beta}_y+\mathbf{w}_{g,1}\mathbf{X}_g\boldsymbol{\delta}_y+\mathbf{w}_{g,1}\mathbf{p}_{gy}\gamma_y+\nu_{g,y})}{\sum_{y'\in\mathcal{Y}}\exp\left(\alpha_{y'}+\mathbf{x}_{g,1}\boldsymbol{\beta}_{y'}+\mathbf{w}_{g,1}\mathbf{X}_g\boldsymbol{\delta}_{y'}+\mathbf{w}_{g,1}\mathbf{p}_{gy'}\gamma_{y'}+\nu_{g,y'}\right)} \\ \vdots \\ \frac{\exp\left(\alpha_y+\mathbf{x}_{g,n_g}\boldsymbol{\beta}_y+\mathbf{w}_{g,n_g}\mathbf{X}_g\boldsymbol{\delta}_y+\mathbf{w}_{g,n_g}\mathbf{p}_{gy}\gamma_y+\nu_{g,y}\right)}{\sum_{y'\in\mathcal{Y}}\exp\left(\alpha_{y'}+\mathbf{x}_{g,n_g}\boldsymbol{\beta}_{y'}+\mathbf{w}_{g,n_g}\mathbf{X}_g\boldsymbol{\delta}_{y'}+\mathbf{w}_{g,n_g}\mathbf{p}_{gy'}\gamma_{y'}+\nu_{g,y'}\right)} \end{pmatrix}. \tag{3}$$

If we collect the L-left hand side vectors from expression (3), we get the $n_g \times L$ matrix $\mathbf{P}_g = (\mathbf{p}_{g0}, \ldots, \mathbf{p}_{gL-1})$. Denoting the right hand side as $\Psi(\cdot)$ we get, where $\boldsymbol{\theta} = \left(\alpha_y, \boldsymbol{\beta}_y, \boldsymbol{\delta}_y, \gamma_y, (\nu_{g,y})_{g\in G}\right)_{y\in\mathcal{Y}}$,

$$\mathbf{P}_g = \Psi(\mathbf{P}_g, \boldsymbol{X}_g, \boldsymbol{W}_g; \boldsymbol{\theta}). \tag{4}$$

Contrary to linear models of social interactions, our framework exhibits multiple equilibria as the structure is compatible with more than one aggregate outcome.[4] Proposition 1 shows that the more alternatives individuals face, the less likely multiple equilibria are. This is due to the independence of errors which implies that with more alternatives, the non–linearities in the fixed point condition become less pronounced, thus enlarging the set of values of $\boldsymbol{\gamma}_y$ for which a unique equilibrium exists.

Proposition 1. ***Multiplicity.*** *In the multinomial choice model with asymmetric influence and network interactions given by (1) and (4), assuming $\gamma_y = \gamma, \delta_y = \delta, \beta_y = \beta, \alpha_y = \alpha, \nu_{g,y} = \nu_g$ for all $y \in \mathcal{Y}$, if $\gamma < 4\left(1 - \frac{1}{L}\right)$ then there is a unique equilibrium.*

[4] The existence of equilibria follows from $A0$, which guarantees that the probability of choosing y is a continuous function bounded by the unit interval, and the application of the Brouwer's fixed theorem (Brock and Durlauf, 2006), as in the standard multinomial logit case.

Review of Economics and Statistics Just Accepted MS.
https://doi.org/10.1162/rest_a_00979

This result is a direct extension to Lee et al. (2014)'s condition for the binary case within an heterogeneous expectations setup (see proof in the Appendix A).[5] Incorporating heterogeneous expectations does not impose more stringent sufficient conditions than the homogeneous one (Brock and Durlauf, 2006).

Due to this multiplicity, we assume that for any given parish, given primitives $(\mathbf{W}_g, \boldsymbol{\theta})$, the data observed by the econometrician is generated from one of the possible equilibria. A unique equilibrium in data condition is common in the literature estimating games with incomplete information using a homogeneous sample of subjects or with spatially similar markets.[6] It is worth highlighting that a unique equilibrium is a sufficient condition for our identification result, detailed in the next section, where we assume that $\mathbf{p}_{g,y}$ is observed by the econometrician.

2.2 Identification

There are three main threats to the identification of the parameters $\boldsymbol{\theta}$. First, in most cases, individuals sort into groups non-randomly. Individuals choose which group they would like to belong to and with whom they would like to interact. If the variables that drive this process of selection are not fully observable, the resulting correlation in unobservables among peers can lead to serious bias in the estimation of social interaction. This self–selection problem

[5] Notice that when $L = 2$ our econometric specification would be equivalent to Lee et al.'s equation (5), with ρ as the endogenous effect. On page 406 of Lee et al.'s paper, the authors state that $|\rho| < 2$ is a sufficient condition for unique equilibrium, which would be equivalent to ours after substituting $L = 2$ on right hand side of the γ inequality.

[6] See De Paula (2013) for a recent discussion on identification and estimation of such multiple equilibria games.

Review of Economics and Statistics Just Accepted MS.
https://doi.org/10.1162/rest_a_00979

can be dealt with by operating under random assignment within *and* across groups based on observables along with group–fixed effects.[7] We use the model specification which has a group–specific component of the error term, and adopt a traditional (pseudo) panel data fixed–effects estimator.

Second, the presence of correlated effects, due to common unobserved information shocks that hit the group as a whole, prevents the separate identification of the exogenous from the endogenous effects in the linear–in–means framework with symmetric influence (Manski, 1993; Brock and Durlauf, 2001; Blume et al., 2010). We show below that even in the presence of correlated effects, it is possible to separately identify the endogenous from the exogenous effects.

Finally, when studying a linear–in–means model, the identification is further complicated by the simultaneity problem, also known as the reflection problem (Manski, 1993). In such models individuals interact in groups, that is individuals are affected by all others in their group and by none outside the group. As a consequence, everyone's behavior affects the others linearly which makes it impossible to separately identify if a group member's action is the cause or the effect of peer's influence. However, the non–linear functional form from the

[7] Assume that agents self–select into different groups in the first step, and that link formation takes place within groups in the second step. Then, as Bramoullé et al. (2009) observe in their conclusion, if link formation is uncorrelated with the observable variables, this two-step model of link formation generates group-fixed effects. This approach is used Lin (2010) with its limitation acknowledged in their footnote 10. Alternative approaches include explicitly correcting for selection (Lee, 1983), imposing exclusion restrictions on the structural model (Graham and Hahn, 2005) or introducing variance restrictions on the error terms that are independent to the group size (Graham, 2008).

Review of Economics and Statistics Just Accepted MS.
https://doi.org/10.1162/rest_a_00979

discrete choice model breaks this simultaneity problem (Brock and Durlauf, 2001, 2006).[8]

Proposition 2. *Under $L > 2$, A.0 and the following assumptions:*

A.1 Joint support of $(\mathbf{x}_{g,i}, \mathbf{w}_{g,i}\mathbf{X}_g)$ is not contained in any linear proper subspace of $\mathbb{R}^{2K}$.

A.2 The support of $\mathbf{w}_{g,i}\mathbf{X}_g$ is not contained in any linear proper subspace of $\mathbb{R}^K$.

A.3 There is a group g such that conditional on $\mathbf{W}_g\mathbf{X}_g$, the support of $\mathbf{x}_{g,i}$ is not contained in any proper linear subspace of $\mathbb{R}^K$.

A.4 There is at least an element k in $\mathbf{x}_{g,i}$, with $\delta_{y,k} \neq 0 \forall y \in \mathcal{Y}$, that does not have bounded support.

A.5 For each y, across different groups, $\mathbf{p}_{g,y}$ and $\nu_{g,y}$ are not constant.

A.6 There is a group g with at least a pair $i, j \in g, i \neq j$, such that $\mathbf{w}_{g,i}\mathbf{p}_{gy} \neq \mathbf{w}_{g,j}\mathbf{p}_{gy}$.

Then, for model described by (1)-(3), the true set of parameters $\boldsymbol{\theta} \setminus (\gamma_y)_{y\in\mathcal{Y}}$ are identified relative to a specific alternative, $y = 0$, while all $(\gamma_y)_{y\in\mathcal{Y}}$ are identified.

Assumption A.0 guarantees the Bayes-Nash equilibrium is given by equation 4. Assumptions A.1-A.3 rule out collinearity between regressors, and assumption A.4 is needed to guarantee that exogenous effects interacted with their corresponding covariates are bounded away from the unit interval which rational expectations belong to. Assumption A.5 imposes

[8] For instance, in the case of symmetric influence multinomial without group unobservables, Blume et al. (2010)'s Theorem 13 provides sufficient conditions for the identification of $\boldsymbol{\theta}$ up to a normalization on one of the alternatives.

Review of Economics and Statistics Just Accepted MS.
https://doi.org/10.1162/rest_a_00979

sufficient between–group variation on expected choices and correlated effects.[9] A.6 requires sufficient within–group variation such that, in at least one group, weighted expected beliefs about other group members should be different for at least two individuals within the group. If these assumptions hold, all endogenous effects are identified and the other structural parameters are identified relative to a specific alternative, even in the presence of correlated effects.

This result is related to the literature that exploits the structure of the weighting matrix (Bramoullé et al., 2009) or group size variations (Lee, 2007) for the identification within a linear–in–means model. We note however, that given the non–linearities at the core of the discrete choice model, the requirement on the structure of the network is weaker than for the continuous case (Bramoullé, 2013). Our result also extends the proof in Brock and Durlauf (2006) that endogenous effects were identified only relative to any distinct alternative.

3 Estimation Procedure and Simulation

Define the *pseudo log-likelihood function*

$$L_N(\boldsymbol{Y} \mid \mathbf{X}, \boldsymbol{W}, \boldsymbol{P}; \boldsymbol{\theta}) = \frac{1}{N}\sum_{g\in G}\sum_{i\in g}\log\left[\frac{\sum\limits_{y\in\mathcal{Y}}\left(\exp\left(\alpha_y + \mathbf{x}_{g,i}\boldsymbol{\beta}_y + \mathbf{w}_{g,i}\mathbf{X}_g\boldsymbol{\delta}_y + \mathbf{w}_{g,i}\mathbf{p}_{gy}\gamma_y + \nu_{g,y}\right)\mathbb{1}_{[y_{g,i}=y]}\right)}{\sum\limits_{y'\in\mathcal{Y}}\exp\left(\alpha_{y'} + \mathbf{x}_{g,i}\boldsymbol{\beta}_{y'} + \mathbf{w}_{g,i}\mathbf{X}_g\boldsymbol{\delta}_{y'} + \mathbf{w}_{g,i}\mathbf{p}_{gy'}\gamma_{y'} + \nu_{g,y'}\right)}\right] \tag{5}$$

[9] Together with the unique equilibrium in data condition, A.5 means that for a given realization of $\mathbf{X}_g$, $\mathbf{W}_g$ and $\boldsymbol{\nu}_g$ only one $\mathbf{P}_g$ should be observed in the data. As we assume that $\boldsymbol{\nu}_g$ is not constant across $g \in G$, even for two groups $g, g' \in G, g \neq g'$ such that $\mathbf{X}_g = \mathbf{X}_{g'}$ and $\mathbf{W}_g = \mathbf{W}_{g'}$ we still get that $\mathbf{P}_g \neq \mathbf{P}_{g'}$, and, therefore, the identification result would still hold when all covariates are discrete.

Review of Economics and Statistics Just Accepted MS.
https://doi.org/10.1162/rest_a_00979

where $N = \sum_{g \in G} n_g$ and $y_{g,i}$ is the occupation chosen by individual $i \in g$. $\boldsymbol{X} \equiv (\boldsymbol{X}_g \forall g \in G)$ denotes the exogenous observables, $\boldsymbol{W} \equiv (\boldsymbol{W}_g \forall g \in G)$ represents the observed weighting matrix, and $\mathbf{P}$ is the collection of $\mathbf{P}_{gy} \forall g, y$.

Estimating the full maximum likelihood estimator (MLE) of our discrete choice problem with social interacions is computationally costly because it is required to repeatedly solve the fixed point of $\mathbf{P} = \boldsymbol{\Psi}(\mathbf{P}, \boldsymbol{X}, \boldsymbol{W}; \theta)$ at each candidate parameter value. Consequently various alternative estimation procedures have been proposed in the literature. We follow Kasahara and Shimotsu (2012, 2018)'s Relaxation Method of the Nested Pseudo Maximum Likelihood (NPL-Λ algorithm) estimation procedure. Let $\hat{\mathbf{P}}^0 \equiv \left(\hat{\mathbf{p}}^0_{gy}\right)_{g \in G, y \in Y}$ be an initial guess of $\mathbf{P}$. Starting from $\hat{\mathbf{P}}^0$, the NPL-Λ algorithm iterates the following steps until $t = T$:

Outer loop: Given $\hat{\mathbf{P}}^{t-1}$, update $\boldsymbol{\theta}$ by $\hat{\boldsymbol{\theta}}^t = \arg\max_{\boldsymbol{\theta} \in \Theta} L_N(\boldsymbol{Y} \mid \mathbf{X}, \boldsymbol{W}, \hat{\mathbf{P}}^{\mathbf{t-1}}; \boldsymbol{\theta})$

Inner loop: Given $\hat{\boldsymbol{\theta}}^t$, $\hat{\mathbf{P}}^t$ solves for the fixed point of $\mathbf{P} = \Lambda\left(\mathbf{P}, \hat{\boldsymbol{\theta}}^t\right) \equiv \boldsymbol{\Psi}\left(\mathbf{P}, \mathbf{X}, \mathbf{W}; \hat{\boldsymbol{\theta}}^t\right)^{\phi} \mathbf{P}^{1-\phi}$ with $\phi \approx 0$.

In the outer loop, we obtain $\hat{\boldsymbol{\theta}}^t$ by maximising the pseudo likelihood function using a Newton–Raphson algorithm. We then solve the fixed point by iterating $\mathbf{P}^j = \Lambda\left(\mathbf{P}^{j-1}, \hat{\boldsymbol{\theta}}^t\right)$ until $\|\mathbf{P}^j - \mathbf{P}^{j-1}\|$ is smaller than a predetermined stopping criterion.[10] This algorithm generates a sequence of estimators $\left\{\hat{\boldsymbol{\theta}}^t, \hat{\mathbf{P}}^t\right\}_{t=1,\ldots,T}$. The estimates at $t = T$ are chosen as the NPL-Λ estimator, which we denote as $\hat{\boldsymbol{\theta}}$.[11]

[10] Bisin et al. (2011) advice implementing such recursive method for $T = 2$ iterations. Following Kasahara and Shimotsu (2018), we fix the tolerance level at 10^{-8} and iterate until $T = 50$. Our main results are based on $\phi = 0.1$. We also estimated with $\phi = 0.8$ and results do not change significantly.

[11] Lee et al. (2014) use a fixed point convergence in their estimation. They substitute the fixed point updating step with $\hat{\mathbf{P}}^t$ being the solution to the fixed point iteration $\mathbf{P} =$

Review of Economics and Statistics Just Accepted MS.
https://doi.org/10.1162/rest_a_00979

The algorithm produces an estimate which converges in probability to the true parameter vector even when the fixed point constraint (4) does not have local contraction properties in a neighborhood of the true parameters. When there is no unobserved heterogeneity or correlated effects, Kasahara and Shimotsu (2012) show in Proposition 5 that even when the belief mapping doesn't satisfy a good contraction property, the NPL-Λ algorithm converges to a consistent estimator provided there is an appropriate value of ϕ (i.e., one that minimizes the spectral radius of $\Lambda(P, \theta^0)$) and a large number of agents.[12] Given our Proposition 1, the contraction property around our true parameters is not guaranteed. This recursive method can be applied to a wide class of dynamic programming models and can account for unobserved heterogeneity.[13]

Allowing for correlated effects at the group level $\nu_{g,y}$ induces an incidental parameters problem which might lead to the inconsistency of maximum likelihood estimators (Neyman and Scott, 1948). This arises from the fact that information about the group fixed–effects

$\Psi(\mathbf{P}, \mathbf{X}, \mathbf{W}; \hat{\boldsymbol{\theta}}^t)$.

[12] The large sample properties of the NPL estimator are proven by Aguirregabiria and Mira (2007) in their Proposition 2.

[13] It is worth noting that there are links to the literature on one-to-one two-sided matching with unobserved heterogeneity. Graham (2013) shows that the equilibrium representation of matches in this model corresponds to the fixed point of a system of nonlinear equations. Mourifié and Siow (2014) introduces peer effects in two-sided matching models and relate it to a multinomial choice with peer effects. In our model, the equilibrium expected occupational choice is also characterized by a fixed point condition which features, as in the one-to-one two-sided matching decision, interdependence between individual actions and (group-level) unobserved heterogeneity.

Review of Economics and Statistics Just Accepted MS.
https://doi.org/10.1162/rest_a_00979

stops accumulating after a finite number of observations as is the case in small groups.[14] In contrast, our application features many groups of large size which allows us to pursue a fixed–effect estimation and circumvent potential specification errors which might arise when using a random-effect approach to modelling unobserved heterogeneity (Dhaene and Jochmans, 2015).[15] In the presence of correlated effects, the requirement for consistency under the NPL-Λ algorithm is that the number of groups and their size go to infinity (Kasahara and Shimotsu, 2012). This is a limitation in some applications with sample data. Gautam (2020) discusses a two-step method for bias reduction in estimation parameters when there is classical measurement error due to sampling within groups.

[14] In a binary choice network model with small groups, the implementation of group fixed–effects is not feasible as it introduces too many fixed–effect parameters to estimate. This is the case in Lee et al. (2014)'s application and they instead follow a random-effect estimation model. Another possible strategy is a conditional maximum likelihood function that differences out the group fixed effects (Andersen, 1970). This approach produces a likelihood function that is not affected by the incidental parameter bias and the estimator converges to the true parameter as the number of groups increases even if they are of small size (Chamberlain, 1980). However, it does not deliver estimates of the fixed effects which are important to recover partial effects (Stammann et al., 2016). Additionally, it does not retain its computational advantage for large-N and large-T.

[15] In Appendix H we perform a Montecarlo simulation to investigate how acute the incidental parameters problem may be for the NPL-Λ algorithm. The results indicate that the larger the number of groups, the larger their size, the closer the our estimates are to the true parameters and the smaller is the dispersion.

Review of Economics and Statistics Just Accepted MS.
https://doi.org/10.1162/rest_a_00979

4 Empirical Application

4.1 Historical Background

Ancient parishes find their origin in the manorial system and remained largely an ecclesiastical unit. Until the seventeenth century, the manor was the principal unit of local administration and justice. However, in due course, parish boundaries came to matter a lot to residents as parishes became public good providers with the "Poor Law" in 1601 giving 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 Metropolis Management Act of 1855 was a landmark in the history of London's government. This Act established the Metropolitan Board of Works and empowered it to develop and implement schemes of London-wide significance. It also created local Board of Works (BW) and Vestries, which were groupings of smaller parishes, with statutory powers to manage and improve local facilities such as paving, lighting, and sewerage. 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 relieved part of the education role which was previously under the control of the established church.

Despite losing importance in terms of civic responsibilities, ecclesiastical parishes remained an integral part of community life. By the end of the nineteenth century Booth (1897) stated "(...) 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 cor-

Review of Economics and Statistics Just Accepted MS.
https://doi.org/10.1162/rest_a_00979

roborated 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). Church attendance was not only mandatory but also important to maintain standing within the community. 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. Parish membership was also important as it determined burial, inclusion in the intentions of the *Missa pro populo*, right to have one's marriage solemnized, etc.

The institutional layout of Victorian London means that both ecclesiastical parishes and BW/Vestries mattered for Londoners. Ecclesiastical parish were important for socialization and would have been a relevant source of information, mentorship, role model and aspiration. Local communities were also relevant sources of job opportunities. Given that parish membership was determined by domicile, we use the ecclesiastical parish boundaries as proxies for social communities. In addition to the religion aspect, this definition of social group also captures the fact that social interactions were very local in nature in late nineteenth century London. The distances over which most people travelled to work remained relatively short. The mean journey to work for those employed in London was only around 5 km in the nineteenth century (Warnes, 1972). We use the spatial distance between individuals as a proxy for the strength of their ties.

Review of Economics and Statistics Just Accepted MS.
https://doi.org/10.1162/rest_a_00979

4.2 Data and Descriptive Statistics

We combine the 100% sample of England and Wales census of 1881 (Schürer and Woollard, 2003) from the North Atlantic Population Project (NAPP, 2015) and digitized historical map of London. The census contains the full address of individuals. In addition to geographic variables, the census also provides a wide range of socio-demographic information. There are over 400 self–reported occupations which we aggregate into three categories: professional, commercial, and industrial (Woollard, 1998).[16] Using historical maps, we geo-reference 5,998 streets of London using points in the middle of each street and overlay ecclesiastical parish and BW/Vestry boundaries. We geographically locate 70% of the entire population in London in the census on the historical map using their place of residence (address, parish and county).

Our sample focuses on native men of working age (between the ages of 15 and 60) who were household heads.[17] This is to ensure that individuals are aware of the institutional

[16] Professional workers include civil service, clerical, legal professions, medical professions, education, liberal arts, scientists and sports. Commercial workers include sales, merchants, dealers of money, and drivers. Industrial workers are further divided into artisan, builder, food/agriculture and services. Industrial food/agriculture include agriculture and food dealers. Workers in industrial services work in service, sales, media and technical sectors. Occupations are broadly defined such that they are not perfect substitutes. There are also domestic occupations and unemployed categories, each representing less than 6% of the male population. We do not include this population in the analysis. This is motivated by the fact that people working in domestic occupations generally live were they work, invalidating our measure of strength of ties.

[17] We remove from our sample foreign-born individuals and individuals who lived in the place where they worked (e.g. prisons, workhouses or other public institutions). Land-owners

Review of Economics and Statistics Just Accepted MS.
https://doi.org/10.1162/rest_a_00979

layout of London and are integrated in a social group. We also restrict ourselves, to avoid small sample issues, to individuals living in parishes for which: the BW/Vestry is composed of at least two ecclesiastical parishes, with at least 30 residents, and with at least one neighbor living on the same street. We end up with a total of 128,709 individuals distributed over 186 ecclesiastical parishes within 32 BW/Vestry. There are on average 5.81 ecclesiastical parishes per BW/Vestry and 692 parishioners per parish. Parish residents have on average 150 neighbors living within 50 meters in the same parish.[18]

Table 1 presents the descriptive statistics of our sample. The mean age was 38. The majority of individuals were married with an average of two children. The average number of servants, which has been used as a proxy for wealth, was 0.175 with a large variation within the sample. Finally 14% of individuals lived in their parish of birth while 47% lived in their county of birth. Figures D.2 and D.3 in the Appendix map these various characteristics and present the spatial clustering of occupations. Professional trades accounted for a large proportion of West London. In contrast, commercial occupations were concentrated in East London. Finally, industrial workers appeared to be more spread out across London.

---

or factory owners often built houses for their workers. We therefore also remove individuals who live at the same address, perform the same job and are not family related.

[18] Figures D.1 and Table D.2 illustrate descriptives of our sample by ecclesiastical parishes

Review of Economics and Statistics Just Accepted MS.
https://doi.org/10.1162/rest_a_00979

Table 1: Summary Statistics per Occupation, London 1881

				Industrial			
	All	Professional	Commercial	Artisan	Builder	Food/Agri	Services
Age	38.221	38.230	37.529	38.817	38.682	38.074	37.872
	(10.40)	(10.16)	(10.31)	(10.64)	(10.35)	(10.29)	(10.37)
Married	0.959	0.910	0.961	0.964	0.971	0.954	0.962
	(0.199)	(0.287)	(0.194)	(0.187)	(0.168)	(0.210)	(0.192)
# children	2.169	1.913	2.040	2.324	2.259	2.055	2.261
	(2.050)	(2.011)	(1.979)	(2.112)	(2.066)	(2.008)	(2.072)
# servants	0.175	0.662	0.106	0.116	0.061	0.339	0.084
	(0.813)	(1.699)	(0.711)	(0.652)	(0.384)	(1.026)	(0.455)
Residents							
Parish birth	0.136	0.059	0.118	0.155	0.164	0.130	0.143
	(0.343)	(0.236)	(0.322)	(0.362)	(0.370)	(0.336)	(0.350)
County birth	0.474	0.323	0.460	0.504	0.497	0.462	0.503
	(0.499)	(0.468)	(0.498)	(0.500)	(0.500)	(0.499)	(0.500)
Obs	128,709	8,371	27,563	30,542	22,065	22,687	17,481

Notes: Mean and standard deviation in parenthesis. Sample includes only native working-age (between 15 and 60) male household heads living in BW/Vestry with at least two ecclesiastical parishes, and living in parish which has a minimum of 30 residents.

Review of Economics and Statistics Just Accepted MS.
https://doi.org/10.1162/rest_a_00979

4.3 Results

4.3.1 Main Results

Following the decision model expressed in equations (2)-(3), each Londoner, taking social group as given, chooses an occupation among professional, commercial or industrial jobs (i.e. artisan, builder, food and agriculture, or services). We control for individual characteristics $\mathbf{x}_{g,i}$ which include age, sex, marital status, number of children, number of servants, and a indicator for whether individuals lived in their parish of birth.[19] Network–level covariates $\mathbf{w}_{g,i}\mathbf{X}_g$ include the same characteristics aggregated at the ecclesiastical parish level and weighted by the geographic distance between individuals of the group (i.e. the strength of ties).[20]

We present the NPL–Λ structural estimation in Table 2. Column 1 shows the results

[19] Our focus is on the role of social contacts on the occupation choice of an individual. We have considered the economic factors as given conditions which impose limits within which these social influences can operate. We restrict our attention to London and including BW/Vestry and parish fixed effects, we are controlling for very localized labor market conditions such as local labor opportunities, working conditions, and schools. However, due to the nature of historical data, human capital variables (such as training, educational qualifications and spells of unemployment), wages and other factors believed to be important in making an occupational choice cannot be measured directly and are therefore excluded from consideration. We might thus be overestimating the effect of social contacts if these factors are positively correlated to social contacts.

[20] In our application, Assumption A.4 in Proposition 2 does not necessarily hold because all exogenous variables have bounded support. Nevertheless, we observe that their support is sufficiently bounded away from the unit interval for at least the following exogenous variables: age, number of servants and number of children.

Review of Economics and Statistics Just Accepted MS.
https://doi.org/10.1162/rest_a_00979

from homogeneous rational expectations à la Brock and Durlauf (2006). That is, every individual is linked to everyone else in the same parish, attaching equal weight to their influence ($w_{g,ij} = 1/(n_g - 1), \forall j \in g \setminus \{i\}$). In column 2, we present our baseline results from the heterogeneous rational expectations model. We define $\mathbf{W}_g$ as row-normalized matrix with zeros on the diagonal and entry $w_{g,ij} = 1/\#nei_{g,i}$ if j is a *neighbor* of i. We define a *neighbor* as any two individuals living within a 50 meters radius from each other's street midpoint within the same parish (i.e., $nei_{g,i} = \{j \in g : j \neq i, d_{g,ij} \leq 50\}$) and $\#nei_{g,i}$ is the number of neighbors i has. In both columns, we control for fixed effects at the BW/Vestry level. Given the role of public good provider of BW/Vestry adding their fixed effect partly deals with the issue of self–selection into groups. Under homogeneous rational expectations, we find that peers within a social group have a positive influence on the occupation choice regardless of the occupation type. However, the degree of influence varies depending on the type of occupation. Peers exert more influence in industrial artisans, industrial builder, and commercial occupations while they exert less influence in the case of industrial services, industrial food/agriculture and professional occupations.[21] Under heterogeneous rational

[21] In the absence of postal codes, we located individuals based on parish of residence but we cannot locate the exact house number. The point geo-referencing a street was located in the middle of the street within the appropriate parish. This means that several individuals living on the same street within the same parish were placed on the same point. This measurement error however occurs only at the street level as individuals are correctly located to their parish of residence. Conley and Topa (2003) discuss the potential implication of imperfect location data on the identification of local interactions. When location information is correct up to some spatial region (in our case ecclesiastical parish), this is equivalent to having only aggregate-level information on the number of people with each outcome within that region.

Review of Economics and Statistics Just Accepted MS.
https://doi.org/10.1162/rest_a_00979

expectations, the influence of peers remains positive for all occupation types. We still observe a large variation in the size of the endogenous effects depending on the type of occupation. However these effects are now more precisely estimated. Therefore adding heterogeneity through the weighting matrix based on geographic distances increases the precision of the estimates.[22]

To further control for potential correlated effects at the group level, we add parish fixed-effects in column 3.[23] These correlated effects can take the form of local industries or an inspiring priest encouraging his parishioners to work or share information. The effects are smaller in magnitude and less significant. This points to the fact that ignoring possible unobserved shocks at the group level might lead us to wrongly attribute effects to the in-

They show that in this case the local identification of the parameters is preserved. The results of such specification is in in column 1 of Table 2.

[22] The inclusion of the heterogeneous rational expectation is particularly appropriate in a setting like ours where members have many peers within a network, reminiscent of a "small world" setup (i.e. a non-sparse adjacency matrix). By "small world" we do not mean that the number of nodes (i.e. number of agents) is small, rather that many of the nodes within the same group are connected to each other. The level of connectedness is also important for smaller communities where we might expect individuals to personally interact with each other and therefore know the characteristics of their peers.

[23] For this specification, we test for the Independence of Irrelevant Alternatives (IIA) assumption in the multinomial logit specification (Hausman and McFadden, 1984). To do so, we take the estimates from the last iteration of the NPL-Λ algorithm. As we have $L = 6$ different occupations, we perform the test by removing each occupation one at a time (apart from the reference occupation which is set to be Professional) and estimating the last step again. In none of the five Hausman-McFadden tests performed do we find evidence to reject the null hypothesis of IIA.

Review of Economics and Statistics Just Accepted MS.
https://doi.org/10.1162/rest_a_00979

fluence of peers. We present the average marginal effects times the standard deviation of the weighted estimated beliefs (AME) associated to the previous specification in column 4.[24] A one standard deviation increase in the expected ratio of peers in a particular occupation leads to an increase in the likelihood of being employed in the same occupation ranging from 0.44% to 3.70% for professionals and industrial artisans respectively.[25] One explanation for the lower influence of peers in professional occupations is that the relevant social contacts for job opportunity is likely to be at a higher level than the ecclesiastical parish. In contrast, industrial artisans, builders and commercial occupations might benefit more from local word-of-mouth social contacts. The use of informal contacts for job search has been found to vary by location and occupation. Our results show that localized social interactions can offer an explanation for differences in occupational choice and the spatial clustering in occupations.[26]

[24] With this scale, we can interpret our reported AME as the effect of a one standard deviation increase in the expected ratio of peers in a particular occupation on the likelihood of being employed in that same occupation. Notice that the standard AME effect simply computes, given parameters $\boldsymbol{\theta}$, the empirical average of the individual change in the estimated propensity to follow an occupation $y \in \mathcal{Y}$, $d\hat{p}_{y,i}$, to a unit change in the expected share of peers following that same occupation, $d\mathbf{w_i}\hat{\mathbf{p}}_y$ ($\tilde{AME}_y = \hat{\mathbb{E}}_i \left[\frac{d\hat{p}_{y,i}}{d\mathbf{w}_i\hat{\mathbf{p}}_y}; \boldsymbol{\theta}\right]$). We scale it by the empirical standard deviation of the expected share of peers following that same occupation ($AME_y = \hat{\mathbb{E}}_i \left[\frac{d\hat{p}_{y,i}}{d\mathbf{w}_i\hat{\mathbf{p}}_y}; \boldsymbol{\theta}\right] \hat{sd}_i \left[\mathbf{w}_i\hat{\mathbf{p}}_y; \boldsymbol{\theta}\right]$).

[25] Notice that the magnitudes found are in line with contemporary studies such as Bayer et al. (2008) who find that two individuals residing on the same versus nearby blocks increase the probability of working together by 0.8% to 3.6%.

[26] In Appendix E.2, we explain and present the average marginal effects of the exogenous effects. Among these exogenous variables, we find that the direct effect is significantly more

Review of Economics and Statistics Just Accepted MS.
https://doi.org/10.1162/rest_a_00979

important than the indirect ones. Age, the number of servants and being married have a negative direct effect on individuals for any occupation category, while the number of children and living in the parish of birth have a positive effect.

Review of Economics and Statistics Just Accepted MS.
https://doi.org/10.1162/rest_a_00979

Table 2: Estimation of endogenous effects, γ_y

	NPL–Λ Expectations			
	Homogeneous	Heterogeneous		AME†
Occupation (y)	(1)	(2)	(3)	(4)
Professional	1.71	1.67**	1.14	0.44%
	(2.84)	(0.68)	(0.77)	[0.002]
Commercial	2.51*	1.96**	2.39***	2.54%
	(1.43)	(0.93)	(0.79)	[0.004]***
Industrial Artisan	3.53***	3.42***	3.14***	3.70%
	(1.27)	(0.33)	(0.55)	[0.004]***
Industrial Builder	2.73	3.32***	3.06***	2.55%
	(2.13)	(0.61)	(0.82)	[0.004]***
Industrial Food/Agriculture	0.73	2.29***	2.93***	2.42%
	(2.46)	(0.85)	(0.63)	[0.003]***
Industrial Services	1.28	2.68*	1.39	0.73%
	(3.53)	(1.57)	(2.66)	[0.008]
log-like	-214,890	-214,250	-212,110	
AIC	430,210	428,940	426,190	
Obs.		128,709		

Review of Economics and Statistics Just Accepted MS.
https://doi.org/10.1162/rest_a_00979

Estimation of endogenous effects, γ_y

	NPL–Λ Expectations			
	Homogeneous	Heterogeneous		AME[†]
Occupation (y)	(1)	(2)	(3)	(4)
Individual characteristics	yes	yes	yes	yes
Group characteristics	yes	yes	yes	yes
BW/Vestry fixed effects	yes	yes	no	no
Parish fixed effects $\nu_{g,y}$	no	no	yes	yes

Notes: Robust standard errors clustered at parish level in parentheses. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. Table reports endogenous effects $\gamma_y, \forall y \in \mathcal{Y}$ from equations (2)-(3) model, where occupation choice set $\mathcal{Y}$ is professional, commercial, industrial artisan, industrial builder, industrial in food and agriculture, or industrial in services. In column 1 every indiviual is linked to everyone else in the same parish, attaching equal weight to their influence ($w_{g,ij} = \frac{1}{(n_g-1)}$). In columns 2 and 3 neighbors are two individuals living within 50 meters with elements of weighting matrix as $w_{g,ij} = \frac{1}{\#nei_{g,i}}$. Individual characteristics x_i include age, marital status, number of children, number of servants and resident in parish of birth. Group characteristics are $\mathbf{w}_{g,i}\mathbf{X}_g$. NPL-$\Lambda$ uses $\phi = 0.1$. [†] Column 4 presents the average marginal effect times the standard deviation of the weighted estimated beliefs (AME). The standard errors in square brackets follow Krinsky and Robb (1986)'s method using $1,000$ replications (see Appendix E for details).

Review of Economics and Statistics Just Accepted MS.
https://doi.org/10.1162/rest_a_00979

4.3.2 Alternative Explanations

Our results show that social networks play a role in the choice of occupation of individuals. There are however other possible mechanisms at play. First, the main threat to our identification is the sorting of individuals into neighborhoods and occupations. We mitigate this problem in our baseline specification by including BW/Vestry fixed effects. We also include parish fixed effects such that social effects are identified by variation within parishes. Nevertheless, one might still worry about within parish-level sorting. Historical evidence points to tight housing market and limited geographic mobility in Victorian London. During the nineteenth century, London grew rapidly due to a high birth rate and migration to the city from other parts of England. This led to severe pressure on the city's housing with many of the inner most districts having over half their population living in overcrowded conditions by the 1880s (Inwood, 2011). There is a large amount of evidence given before the Royal Commission on Means of Locomotion and Transport in London which suggests that there was a large unsatisfied demand for housing in London (Habakkuk, 1962).

We also examine whether the decision to move away from the parish of birth and residing on a particular street-parish in 1881 is correlated with the composition of occupations present on that street-parish in 1851.[27] In Table F.4 in the Appendix we see that the share of individuals working in different occupation types in 1851 at the street of residence in 1881 do not influence the decision to move away from the parish of birth in 1881. Variables that do matter are the average number of servants and the average share of residents that were

[27] We first match parish and street addresses from the I-CeM 1851 census (see Appendix B.3) to our baseline sample from the NAPP 1881 census. We then aggregate characteristics of the population in 1851 at the street-parish level.

Review of Economics and Statistics Just Accepted MS.
https://doi.org/10.1162/rest_a_00979

born in that parish. This suggests that there is no sorting at the street-parish level based on historical expertise.

We further dispel concerns about within-parish sorting in Table F.5. We first examine whether there are differences in the endogenous effects by mobility status. We interact the various occupation types with an indicator of whether the individual lives in his parish of birth (i.e. non-movers, in column 2). We do not see an additional effect for non-movers, except for the industrial service occupations. We also implement two placebo tests. In the first one, we randomly allocate individuals to different streets *within* the same parish. The "placebo neigbhors" consists of parish members who have been allocated to a street within a 50 meter radius of an individual but who in reality may live further away. We do so 100 times and estimate our structural parameters for each new allocation (column 3). In the second one, we define an individual's neighbors as parish members who are more distant than the 50 meter threshold (column 4). In the first placebo, the endogenous effects are all smaller in magnitude than the benchmark. In the second one, they are statistically different to the benchmark. With the exception of one occupation, all others are not significantly different than zero. Given the difference in the identity of neighbors, it is not surprising that we cannot recover our baseline results with these placebos. This reveals the importance of correctly identifying neighbors and their the geographic location within a parish. Specifically, the geographic proximity between individuals is a sensible proxy for the likelihood and frequency of their encounters and consequently the influence they have on one another.

Second, Victorian Britain has been depicted as an unequal and immobile society. Family background has been found to exert greater influence on economic status than was previously believed (Solon et al., 1991). If individuals inherited their occupations from their family and

Review of Economics and Statistics Just Accepted MS.
https://doi.org/10.1162/rest_a_00979

simultaneously chose to live close to their parents, we might be attributing occupational choice to social network instead of family. To control for the role of family, we identify the occupation of fathers and brothers in our sample and control for their potential influence in the occupational choice. We use the 100% 1851 and 1881 datasets from I-CeM Project (Schürer and Higgs, 2020) and link individuals using a novel linking method (Abramitzky et al., 2020) (see Appendix C for details on the linking procedure).[28] This allows us to track individuals between the 1851 and 1881 censuses. We are thus able to identify the father as the head of household in 1851 and the son as the head of household himself in 1881. We identify also their brothers living in the same household in 1851 and track them in the 1881 census. Table F.6 in the Appendix compares the fathers' occupations observed in 1851 and sons' occupations observed in 1881. To take into account the age of occupational choice, the sample is based on sons aged 13-19 in 1851 when their father's occupation is observed. The occupations of sons are observed in 1881. We see that 63% of sons ended up in occupations different from those of their fathers. This confirms the conclusion reached by Long (2013) that the intergenerational mobility was larger than previously believed.

In Table F.7 in the Appendix we control for family effects in two ways. At the last step of our recursive baseline estimation, we first control for the father's occupation for sample of linked individuals whose father's occupation is known. Comparing columns 1 and 2, we see that controlling for the father's occupation does not change the sign, magnitude nor significance of the endogenous effects. We then control for the brother's occupation for the

[28] Our main results are based on the NAPP 1881 census. Since starting the project, the NAPP no longer provides the names and addresses of individuals. We therefore cannot rely on the I-CeM project for the linking of the 1851 to 1881 censuses.

Review of Economics and Statistics Just Accepted MS.
https://doi.org/10.1162/rest_a_00979

sample of linked individuals who have brothers with known occupation. Again, comparing columns 3 and 4, the coefficients of the endogenous effects are very similar with and without the brother's occupation as a control. These results suggest that social contact play a significant role even after controlling for the role of family members.

4.3.3 Robustness Checks

We perform a number of robustness checks. First, we explore the sensitivity of our results to different specifications of the weighting matrix as in Lin (2010). In column 2 of Table G.8 in the Appendix, neighbors are defined as individuals living on the same street.[29] Results are very similar to our main specification in column 1, where neighbors are living 50 meters from one another in the same parish. In column 3 weights are created based on the continuous distance between parish members. We use the row-normalized exponential distance weights which takes the form $w_{g,ij} = \frac{\exp(-\alpha d_{g,ij})}{\sum_{j' \in g \setminus \{i\}} \exp(-\alpha d_{g,ij'})}$, $\forall j \in g \setminus \{i\}$ where α is any positive exponent (we set it equal to 1). Closer parish members exert more influence than members who live further away. This reflects the probability or frequency of meeting and hence the strength of ties. Again, the endogenous effects are robust to this alternative measure of tie strength.

Second, in the baseline specification, working-age men form rational expectations about the occupational choice of all other working-age men within their parish. The age of individuals and their peers are included as control characteristics. However older and experienced workers have already chosen their occupation and will no longer form expectations. We

[29] Notice that our parameters are identified provided that, within a parish, streets have different numbers of residents, thus guaranteeing that the rows of the weighting matrix differ for two residents living on different streets.

Review of Economics and Statistics Just Accepted MS.
https://doi.org/10.1162/rest_a_00979

therefore restrict our sample to younger individuals who form beliefs on the occupational choice of other parish members from their cohort or younger, and use the actual occupation choice of parish members older than them. This allows us to include all parish members in the interaction matrix even though we restrict the sample to younger individuals. Table G.9 presents the baseline in column 1, restricts the sample to men between 15 and 30 years old in column 2, and restricts the sample to men between 15 and 45 years old in column 3. As we restrict to younger individuals, the endogenous coefficients are larger in magnitude (except for industrial food/agriculture) and significant at smaller levels than in the baseline specification. This can be explained by the fact that we are only partly capturing the social group of older cohorts as individuals may have moved or passed away.

Third, as stated in proposition 1, multiple equilibria could arise whenever endogenous effects are larger than 3.33 in absolute value. We investigate whether there are multiple equilibria in two different exercises using first our estimated parameters for $\boldsymbol{\gamma}$ and then these estimated parameters but adding one time the standard error. None of the 186 parishes included in our study have more than one equilibrium in both exercises.[30]

Finally, to remove the first-order bias term in our NPL-Λ estimates stemming from the incidental parameter bias, we define a split-panel jackknife estimation procedure based on unbalanced panel data models (see Appendix I). Taken together, the results suggest that if any, the incidental parameters bias is not acute in our application.

[30] We implemented a spectral method and results are available upon request. We use 1,000 randomly generated starting values where all subjects follow each occupation with strictly positive probability. In addition, we set 6 "extreme" starting values, where every parish member follows each occupation with certainty.

Review of Economics and Statistics Just Accepted MS.
https://doi.org/10.1162/rest_a_00979

5 Conclusion

This paper investigates how peers within a social group influence the occupational choice of individuals in Victorian London. We propose a multinomial choice model with social interactions and asymmetric influences, thereby extending prior work on binary choice models with asymmetric influences (Lee et al., 2014) and multinomial choice models with symmetric influences (Brock and Durlauf, 2006). The model allows for correlated effects at the group level and includes a spatial weighting matrix to capture the strength of social ties. We establish the identification of the endogenous and exogenous effects when there is enough variation in the weighting matrix and provide a sufficient condition for a unique equilibrium.

We construct a new dataset which geographically locates London residents enumerated in the full census of 1881. The boundaries of ecclesiastical parishes are used as a proxy for social groups while the strength of the ties between parish members is measured by their geographical distance. To circumvent the self-selection into social group problems, we use fixed–effects at the administrative and social group level. The Relaxation Method Nested Pseudo Likelihood algorithm proposed by Kasahara and Shimotsu (2012) is applied to provide consistent and asymptotically efficient estimates of the structural parameters.

Our results indicate that social interactions within parishes played an important role in determining occupational choices. We find robust empirical evidence that an increase in the share of peers within a parish employed in a particular occupation led to a significant increase an individual's probability of being employed at that same occupation. The magnitude of the effects ranges from 0.44% to 3.70% depending on the type of occupation. Individuals working as industrial artisans are the most influenced by the peers within their social groups

Review of Economics and Statistics Just Accepted MS.
https://doi.org/10.1162/rest_a_00979

while individuals in professional occupations are the least influenced.

While our results rely on historical data, our setup might offer a more credible definition of social network than contemporary studies. In the modern world of easy mobility and technological information, geographical–based measures may not adequately capture social networks. In contrast, communities in the nineteenth century were arguable more local in nature. Moreover, the religious feature of our social group definition offers a plausible additional dimension given that church attendance remained mandatory. More generally, the empirical model may be applied to many areas involving local interactions and categorical outcomes. We show that failing to account for asymmetric influence may bias the endogenous effect on occupational choices.

These findings contribute to our understanding of spatial clustering in occupation and inequality patterns within a city (Bayer et al., 2008; Glaeser et al., 2009). It also provides a novel interpretation for high intergenerational persistence of segregation by occupation that has been documented by Borjas (1994) if individuals use their inherited social connections to find jobs.

References

Abramitzky, Ran, Roy Mill, and Santiago Pérez. “Linking individuals across historical sources: A fully automated approach.” *Historical Methods: A Journal of Quantitative and Interdisciplinary History* 53 (2020): 94–111.

Aguirregabiria, Victor and Pedro Mira. “Sequential estimation of dynamic discrete games.” *Econometrica* 75 (2007): 1–53.

Review of Economics and Statistics Just Accepted MS.
https://doi.org/10.1162/rest_a_00979

Albornoz, Facundo, Antonio Cabrales, and Esther Hauk. "Occupational choice with endogenous spillovers." *The Economic Journal* 129 (2017): 1953–1970.

Andersen, Erling Bernhard. "Asymptotic properties of conditional maximum-likelihood estimators." *Journal of the Royal Statistical Society. Series B (Methodological)* 32 (1970): 283–301.

Banerjee, Abhijit V and Andrew F Newman. "Occupational choice and the process of development." *Journal of Political Economy* 101 (1993): 274–298.

Bayer, Patrick, Stephen Ross, and Giorgio Topa. "Place of Work and Place of Residence: Informal Hiring Networks and Labor Market Outcomes." *Journal of Political Economy* 116 (2008): 1150–1196.

Bisin, Alberto, Andrea Moro, and Giorgio Topa. "The empirical content of models with multiple equilibria in economies with social interactions." *National Bureau of Economic Research No. 17196* (2011).

Blume, Lawrence E, William A Brock, Steven N Durlauf, and Yannis M Ioannides. "18 Identification of social interactions." *Handbook of Social Economics.* Volume 1B . Cambridge: MIT Press. 2010. 853–964.

Blume, Lawrence E, William A Brock, Steven N Durlauf, and Rajshri Jayaraman. "Linear social interactions models." *Journal of Political Economy* 123 (2015): 444–496.

Booth, Charles. *Life and Labour of the People of London.* Manchester University Press, 1897.

Review of Economics and Statistics Just Accepted MS.
https://doi.org/10.1162/rest_a_00979

Borjas, George J. "Immigration, ethnic identity, and assimilation: The intergenerational transmission of immigrant skills." *Economic aspects of international migration.* . Springer, 1994. 139–154.

Bramoullé, Yann. "Comment to Social networks and the identification of peer effects." *Journal of Business & Economic Statistics* 31 (2013): 264–266.

Bramoullé, Yann, Habiba Djebbari, and Bernard Fortin. "Identification of peer effects through social networks." *Journal of Econometrics* 150 (2009): 41–55.

Brock, William A and Steven N Durlauf. "Discrete choice with social interactions." *The Review of Economic Studies* 68 (2001): 235–260.

Brock, William A and Steven N Durlauf. "A Multinomial Choice Model with Neighborhood effects." *American Economic Review* 92 (2002): 298–303.

Brock, William A and Steven N Durlauf. "Multinomial Choice with Social Interactions." *The economy as an evolving complex system.* Ed. L.E. Blume and S.N. Durlauf. Volume 3 . Oxford University Press, 2006.

Chamberlain, Gary. "Analysis of Covariance With Qualitative Data." *The Review of Economic Studies* 47 (1980): 225–238.

Cicala, Steve, Roland G Fryer, and Jörg L Spenkuch. "Self-selection and comparative advantage in social interactions." *Journal of the European Economic Association* 16 (2017): 983–1020.

Review of Economics and Statistics Just Accepted MS.
https://doi.org/10.1162/rest_a_00979

Conley, Timothy G and Giorgio Topa. "Identification of local interaction models with imperfect location data." *Journal of Applied Econometrics* 18 (2003): 605–618.

De Paula, Aureo. "Econometric analysis of games with multiple equilibria." *Annual Review of Economics* 5 (2013): 107–131.

Dhaene, Geert and Koen Jochmans. "Split-panel jackknife estimation of fixed-effect models." *The Review of Economic Studies* 82 (2015): 991–1030.

Friedman, Benjamin M. "Economics: A moral inquiry with religious origins." *American Economic Review* 101 (2011): 166–170.

Gautam, Sanghmitra. "Quantifying welfare effects in the presence of externalites: An ex-ante evaluation of a sanitation intervention." *Mimeo* (2020).

Glaeser, Edward L, Matt Resseger, and Kristina Tobio. "Inequality in cities." *Journal of Regional Science* 49 (2009): 617–646.

Graham, Bryan. "Identifying social interactions through conditional variance restrictions." *Econometrica* 76 (2008): 643–660.

Graham, Bryan. "Comparative static and computational methods for an empirical one-to-one transferable utility matching model." *Structural Econometric Models*. . Emerald Group Publishing Limited, 2013. 153–181.

Graham, Bryan and Jinyong Hahn. "Identification and estimation of the linear-in-means model of social interactions." *Economics Letters* 88 (2005): 1–6.

Review of Economics and Statistics Just Accepted MS.
https://doi.org/10.1162/rest_a_00979

Habakkuk, Hrothgar. "Fluctuations in house-building in Britain and the United States in the Nineteenth Century." *The Journal of Economic History* 22 (1962): 198–230.

Hausman, Jerry A and Daniel McFadden. "Specification tests for the multinomial logit model." *Econometrica* 5 (1984): 1219–1240.

Inwood, Stephen. *City of cities: The birth of modern London*. Pan Macmillan, 2011.

Ioannides, Yannis and Linda Loury. "Job information networks, neighborhood effects, and inequality." *Journal of Economic Literature* 42 (2004): 1056–1093.

Kasahara, Hiroyuki and Katsumi Shimotsu. "Sequential Estimation of Structural Models with a Fixed Point Constraint." *Econometrica* 80 (2012): 2303–2319.

Kasahara, Hiroyuki and Katsumi Shimotsu. "Estimation of discrete choice dynamic programming models." *The Japanese Economic Review* 69 (2018): 28–58.

Krinsky, Itzhak and Leslie Robb. "On approximating the statistical properties of elasticities." *The Review of Economics and Statistics* 68 (1986): 715–719.

Lee, Lung-Fei. "Generalized econometric models with selectivity." *Econometrica* 51 (1983): 507–512.

Lee, Lung-fei. "Identification and estimation of econometric models with group interactions, contextual factors and fixed effects." *Journal of Econometrics* 140 (2007): 333–374.

Lee, Lung-fei, Ji Li, and Xu Lin. "Binary choice models with social interactions under heterogeneous rational expectations." *The Review of Economic and Statistics* 96 (2014): 402–417.

Review of Economics and Statistics Just Accepted MS.
https://doi.org/10.1162/rest_a_00979

Lin, Xu. "Identifying peer effects in student academic achievement by spatial autoregressive models with group unobservables." *Journal of Labor Economics* 28 (2010): 825–860.

Long, Jason. "The surprising social mobility of Victorian Britain." *European Review of Economic History* 17 (2013): 1–23.

Manski, Charles F. "Identification of endogenous social effects: The reflection problem." *The Review of Economic Studies* 60 (1993): 531–542.

McFadden, Daniel L. *Conditional Logit Analysis of Qualitative Choice Behavior.* in Paul Zarembka, ed. Frontiers in Econometrics (New York: Academic Press), 1974.

Mourifié, Ismael and Aloysius Siow. "Cohabitation versus marriage: Marriage matching with peer effects." *University of Toronto Mimeo* (2014).

Munshi, Kaivan. "Networks in the modern economy: Mexican migrants in the US labor market." *The Quarterly Journal of Economics* 118 (2003): 549.

Murphy, Kevin M, Andrei Shleifer, and Robert W Vishny. "The allocation of talent: Implications for growth." *The Quarterly Journal of Economics* 106 (1991): 503–530.

NAPP. "Minnesota Population Center. North Atlantic Population Project: Complete Count Microdata. Version 2.2 [Machine-readable database]." *Minneapolis: Minnesota Population Center* (2015).

Neyman, Jerzy and Elizabeth L Scott. "Consistent estimates based on partially consistent observations." *Econometrica* 16 (1948): 1–32.

Review of Economics and Statistics Just Accepted MS.
https://doi.org/10.1162/rest_a_00979

Patacchini, Eleonora and Yves Zenou. "Ethnic networks and employment outcomes." *Regional Science and Urban Economics* 42 (2012): 938–949.

Patel, Krishna and Francis Vella. "Immigrant networks and their implications for occupational choice and wages." *The Review of Economics and Statistics* 95 (2013): 1249–1277.

Schürer, Kevin and Edward Higgs. Integrated Census Microdata (I-CeM), Names and Addresses, 1851-1911: Special Licence Access. [data collection]. 2nd Edition. UK Data Sevice, SN:7856, http://doi.org/10.5255/UKDA-SN-7856-2, apr 2020.

Schürer, Kevin and Matthew Woollard. "National Sample from the 1881 Census of Great Britain." *Colchester, Essex: History Data Service, UK Data Archive [distributor]* (2003).

Smith, Richard Mudie. *The religious life of London*. Hodder and Stoughton, 1904.

Soetevent, Adriaan R and Peter Kooreman. "A discrete-choice model with social interactions with an application to high school teen behavior." *Journal of Applied Econometrics* 22 (2007): 599–624.

Solon, Gary. "Intergenerational mobility in the labor market." *Handbook of Labor Economics*. . Volume 3 . Elsevier, 1999. 1761–1800.

Solon, Gary, Mary Corcoran, Roger Gordon, and Deborah Laren. "A Longitudinal Analysis of Sibling Correlations in Economic Status." *Journal of Human Resources* 26 (1991).

Stammann, Amrei, Florian Heiss, and Daniel McFadden. "Estimating fixed effects logit models with large panel data." *Mimeo* (2016).

Review of Economics and Statistics Just Accepted MS.
https://doi.org/10.1162/rest_a_00979

Warnes, Anthony. "Estimates of journey-to-work distances from census statistics." *Regional Studies* 6 (1972): 315–326.

Woollard, Matthew. "The classification of occupations in the 1881 census of England and Wales." *History and Computing* 10 (1998): 17–36.