No.47

JULIO DE 2017

Documentos **CEDE**

ISSN 1657-7191 Edición electrónica.

Multinomial choice with social
interactions: occupations in
Victorian London

José-Alberto Guerra
Myra Mohnen

CEDE
CENTRO DE ESTUDIOS SOBRE DESARROLLO ECONÓMICO

Universidad de
los Andes
Facultad de Economía

Serie Documentos Cede, 2017-47
ISSN 1657-7191 Edición electrónica.
Julio de 2017

© 2017, Universidad de los Andes, Facultad de Economía,
CEDE. Calle 19A No. 1 – 37 Este, Bloque W.
Bogotá, D. C., Colombia Teléfonos: 3394949- 3394999,
extensiones 2400, 2049, 2467
infocede@uniandes.edu.co
<http://economia.uniandes.edu.co>

Impreso en Colombia – Printed in Colombia

La serie de Documentos de Trabajo CEDE se circula con propósitos de discusión y divulgación. Los artículos no han sido evaluados por pares ni sujetos a ningún tipo de evaluación formal por parte del equipo de trabajo del CEDE.

El contenido de la presente publicación se encuentra protegido por las normas internacionales y nacionales vigentes sobre propiedad intelectual, por tanto su utilización, reproducción, comunicación pública, transformación, distribución, alquiler, préstamo público e importación, total o parcial, en todo o en parte, en formato impreso, digital o en cualquier formato conocido o por conocer, se encuentran prohibidos, y sólo serán lícitos en la medida en que se cuente con la autorización previa y expresa por escrito del autor o titular. Las limitaciones y excepciones al Derecho de Autor, sólo serán aplicables en la medida en que se den dentro de los denominados Usos Honrados (Fair use), estén previa y expresamente establecidas, no causen un grave e injustificado perjuicio a los intereses legítimos del autor o titular, y no atenten contra la normal explotación de la obra.

Universidad de los Andes | Vigilada Mineducación
Reconocimiento como Universidad: Decreto 1297 del 30 de mayo de 1964. Reconocimiento personería jurídica: Resolución 28 del 23 de febrero de 1949 Minjusticia.

Multinomial choice with social interactions: occupations in Victorian London^{*}

José-Alberto Guerra[†], Myra Mohnen[‡]

Abstract

This paper presents a multinomial choice model with social interactions in an incomplete network. Individuals form heterogeneous rational expectations about the behavior of peers by taking into account their characteristics and the strength of their ties. We show the network conditions under which the endogenous and exogenous effects can be separately identified even in the presence of correlated effects at the group level. Conditions for unique equilibrium are established. We apply our empirical model to occupational choice in nineteenth century London. Thanks to a newly constructed dataset, we use ecclesiastical parish boundaries as proxies for social groups and geographic distances between individuals as measures of the strength of their ties. Our results show that endogenous network effects were important above and beyond correlated and exogenous effects. We uncover distinct impact by occupation type: peers in professional and industrial occupations have a positive impact on the likelihood of following a similar occupation while commercial have a negative one.

Keywords: Social interactions, Heterogeneous beliefs, Multinomial choice, Occupational choice
JEL codes: C25, C31, J24, N93

^{*}We are grateful to Orazio Attanasio, Patrick Bayer, Sandra Black, Lawrence Blume, Yann Bramoullé, Antonio Cabrales, Syngjoo Choi, Rachel Kranton, Aureo de Paula, Imran Rasul, Uta Schöenberg, Martin Weidner, Yves Zenou and audiences at UCL, IFS, EDePo, Cambridge-INET Institute, HCEO Chicago Summer School in Social Economics, ITAM, Universidad del Rosario, University of Melbourne, Universidad de los Andes for valuable 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.

[†] *corresponding author*, ja.guerra@uniandes.edu.co, [‡] myra.mohnen@essex.ac.uk

Elección multinomial con interacciones sociales: Decisión ocupacional en el Londres Victoriano[★]

José-Alberto Guerra[†], Myra Mohnen[‡]

Abstract

Este artículo presenta un modelo de elección multinomial con interacciones sociales en una red incompleta. Los individuos forman expectativas racionales heterogéneas sobre el comportamiento de sus pares teniendo en cuenta sus características y la fuerza de sus lazos. Se muestran las condiciones de la red bajo las cuales los efectos endógenos y exógenos pueden identificarse por separado incluso en presencia de efectos correlacionados a nivel de grupo. Se establecen las condiciones para un equilibrio único. Aplicamos nuestro modelo empírico a la elección ocupacional en el Londres del siglo XIX. Gracias a un conjunto de datos históricos, usamos las fronteras de las parroquias eclesiásticas como aproximación a los grupos sociales y la distancia geográfica entre individuos como medida de la fuerza de sus lazos. Nuestros resultados muestran que los efectos endógenos son importantes más allá de los efectos correlacionados a nivel de grupo. Encontramos que hay efectos distintos según el tipo de ocupación: el efecto endógeno de grupo par es positivo sobre la probabilidad de seguir una ocupación industrial o profesional mientras que es negativo para las ocupaciones comerciales.

Keywords: Interacciones sociales, Expectativas heterogéneas, Elección multinomial, Decisión ocupacional

Clasificación JEL: C25, C31, J24, N93

[★]Agradecemos a Orazio Attanasio, Patrick Bayer, Sandra Black, Lawrence Blume, Yann Bramoullé, Antonio Cabrales, Syngjoo Choi, Rachel Kranton, Aureo de Paula, Imran Rasul, Uta Schöenberg, Martin Weidner, Yves Zenou y audiencias en UCL, IFS, EDePo, Cambridge-INET Institute, HCEO Chicago Summer School in Social Economics, ITAM, Universidad del Rosario, University of Melbourne, Universidad de los Andes por comentarios valiosos. Agradecemos al Departamento de Arqueología del Museo de Londres por facilitarnos el Mapa Histórico de Londres y a Daniel Felipe Martínez Enriquez por su asistencia en la extensión de este.

[†] *autor de contacto*, ja.guerra@uniandes.edu.co, [‡] myra.mohnen@essex.ac.uk

1. Introduction

It is widely acknowledged that individuals are embedded in a social network and their behaviors and outcomes are influenced by these social interactions.¹ In a seminal paper, Manski (1993) proposes a linear-in-means model with social interactions within a complete network. In this framework, individuals belong to a group, interact with all the other members of the group (i.e. their peers), and form rational expectations regarding their peers' choice probabilities. Consequently, individuals are equally influenced by all members of their own group and form homogenous rational expectations based on the group-level behavior (i.e. the mean of the group). This homogenous rational expectation framework is adopted by Brock and Durlauf (2001) in a discrete choice model and Brock and Durlauf (2002, 2006) in a multinomial choice model. However, in the case of an incomplete network where individuals do not all interact with each other but only with a small group within the network, individuals may take into account their peers' characteristics thereby forming heterogenous rational expectations. Lee et al. (2014) provide such an extension within a binary choice model.

In this paper, we present a multinomial choice model with heterogeneous beliefs thus providing an extension to Lee et al. (2014)'s binary choice with heterogeneous beliefs and Brock and Durlauf (2002, 2006)'s multinomial choice with homogeneous beliefs. Specifically, individuals interact with members of a local group within a larger social network and their decision is influenced by their peers. When making a decision, individuals form rational expectations about their peers' decisions taking into account their characteristics.² We introduce heterogeneity in these rational expectations through a weighting matrix which captures the strength of ties between peers. We allow for correlated effects at the group level which capture potential shocks hitting the group as a whole.

We explore the effect of the social network structure on identification. We show that the asymmetry in the influence of one's peers, introduced by the weighting matrix, allows us to separately identify the endogenous effect (the influence of peers' outcomes) and the exogenous effect (the influence of peers' attributes) from group unobservables. We establish the conditions under which a unique equilibrium can be found. We estimate the model using a Recursive Pseudo Maximum Likelihood (PML) with an equilibrium fixed point subroutine following the Relaxation Method proposed by Kasahara and Shimotsu (2012). This method converges to the true parameters even when the fixed point constraint does not have local contraction property in the neighborhood of the true parameters. In light of our Monte Carlo simulations, we confirm that whenever the number of individuals per group is large, the incidental parameters problem is not as severe as in short panel case.³

The empirical framework developed in this paper may be applied to many areas involving local interactions and categorical outcomes such as criminal activities, modes of transport, or

¹The importance of peer effects have been applied to topics as diverse as criminal activity (e.g., Glaeser et al. (1996)), welfare participation (Bertrand et al. (2000)), post-war employment of World War I veterans (e.g., Laschever (2009)), school achievement (e.g., Sacerdote (2001); De Giorgi et al. (2007)), participation in retirement plans (Duflo and Saez (2003)), and obesity (e.g., Trogon et al. (2008)).

²We look at small groups setting where individuals form rational expectations on their behaviors instead of directly observing their behavior as in Soeteven and Kooreman (2007) and Graham and Hahn (2005) or modeled via a game-theoretical framework as in Krauth (2006). Aradillas-Lopez (2010) considers an incomplete information binary model in a semiparametric setting.

³Reminiscent of the non-linear panel data with fixed effects literature (Arellano and Bonhomme, 2011). It also depicts how variation across and within groups may be exploited for identification with non-network data (Bramoullé et al., 2009; Bramoullé, 2013; Goldsmith-Pinkham and Imbens, 2013).

technology adoption. We consider individuals' occupational choice under the influence of their peers in nineteenth century London. By constructing a new dataset combining a geo-referenced historical map of London and the 1881 100% census, we can locate individuals down to the street level. Given that parishes played an important social role in the community and that parish membership were based on residency, parish boundaries are used as a proxy for social groups. The geographical distances between households are used as inputs for the weighting matrix.

Our results highlight the importance of social networks on occupational choice. We show that failing to account for asymmetric influences and ignoring possible correlated effects may bias endogenous effects on occupational choices upwards. We look at both endogenous and exogenous effects, and uncover distinct endogenous effects on labor outcomes depending on the type of occupation. Peers in professional and industrial occupations have a positive effect on the likelihood of following a similar occupation while peers in commercial occupations in contrast have a negative effect.

In addition to the growing literature on multinomial choice models with social interactions, this paper also contributes to the literature interested in measuring the empirical relevance of social networks on labor market outcomes (Topa, 2011). Prior empirical work on the effect of social contacts on employment status and unemployment duration generally confirms that contacts are individually beneficial to workers (Cappellari and Tatsiramos, 2015; Blau and Robins, 1990; Kramarz and Skans, 2007; Akerlof and Kranton, 2000). Credibly identifying social interactions remains challenging. First, researchers usually proxy the relevant group using some arbitrary metric of distance based on social and/or geographical proximity such as school (Gaviria and Raphael, 2001; Hoxby, 2000), grade (Hanushek et al., 2003), rooms and dorms (Sacerdote, 2001), neighboring families (Solon et al., 2000; Bayer et al., 2008; Helmers and Patnam, 2014), and gender or race (Patacchini and Zenou, 2012). In the modern world of high mobility and easy access to communication technologies, there is a legitimate concern that physical distances nowadays may have become less important in shaping social networks. Focusing on a historical time period where social networks were mostly geographically localized, our empirical application has the advantage of providing us with a more credible proxy for social groups. Moreover, the religious feature of parishes offers an additional relevant dimension to social groups in a time period where religion played a central role (Booth, 1897). Second, we include parish-level fixed effects to control for the possible confounding effects in the form of local industry shocks or a local priest who might encourage parishioners to work or share information. To partially deal with self-selection concerns, we control for local institutions which might factor in the location decision of individuals (Bayer and Timmins, 2007).

The rest of the paper is organized as follows. Section 2 presents the multinomial choice model with social interactions and heterogeneous beliefs. We characterize the set of parameters that are sufficient for a unique equilibrium and show that our structural parameters are identified. Section 3 shows the estimation strategy together with the results of the Monte Carlo simulation. To illustrate our model, section 4 implements our approach to occupational choices in nineteenth century London. We first explain the social role of parishes to justify the use of parish boundaries as a social group definition. After summarizing the data, we present the results of the endogenous effects on the occupational choice of individuals. We finally perform several robustness checks. In section 5 we summarize our findings and conclude.

2. Empirical Model

Consider individual i embedded in a group g faced with a multinomial choice $y \in Y = \{0, 1, \dots, L-1\}$. There are G groups in the population, each with n_g members. The structure of group $g \in G$ determines the social interactions between individuals and consequently the extent of their influence on each other's decision. It is defined by an adjacency matrix $\mathbf{W}_g$, with entry $w_{g,ij} \forall i, j \in g$, representing the strength of the tie between i and j . We assume self-influence is excluded ($w_{g,ii} = 0$). The $1 \times n_g$ vector $\mathbf{w}_{g,i}$ determines the weight individual i , belonging to group g , places on each member of her group.⁴ In particular, i can only be influenced by her neighbors $nei_{g,i} = \{j \in g \mid w_{g,ij} \neq 0\}$.⁵

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 $y \in Y$ such as to maximize a pay-off function $V : Y \rightarrow \mathbb{R}$, 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 $v_{g,y}$, and a random private utility $\epsilon_{y,i}$ (Brock and Durlauf, 2006; Soeteven and Kooreman, 2007).

$$\max_{y \in Y} u(y, \mathbf{x}_i) + S(y, \mathbf{X}_g, \mathbf{p}_{gy}^e) + v_{g,y} + \epsilon_{y,i}, \quad (1)$$

where $\mathbf{X}_g$ is the $n_g \times K$ matrix with i^{th} -row equal to $\mathbf{x}_{g,i}$. We assume that $\mathbf{W}_g$, $\mathbf{X}_g$, $v_{g,y}$ and $\epsilon_{y,i}$ are in individual i 's information set ($\mathbb{I}_i$). Individual i forms rational expectations regarding the outcome decision of members of her group ($\mathbf{p}_{gy}^e$)⁶ while taking into account their specific characteristics ($\mathbf{X}_g$) and the strength of the social ties linking them ($\mathbf{W}_g$). $v_{g,y}$ and $\epsilon_{y,i}$ are unobservable to the econometrician and independent across outcomes.

We assume that equation 1 has a linear functional form and $\epsilon_{y,i}$ are independent and identically distributed across and within groups g with known Gumbel distribution function $F_\epsilon(\epsilon_{y,i} < \epsilon) = \exp(-\exp(\epsilon))$.⁷ Consequently, agent i , belonging to social group g , chooses y with probability given by⁸

$$p_{gy,i} = \frac{\exp(\alpha_y + \mathbf{x}_i \beta_y + \mathbf{w}_{g,i} \mathbf{X}_g \delta_y + \mathbf{w}_{g,i} \mathbf{p}_{gy}^e \gamma_y + v_{g,y})}{\sum_{y' \in Y} \exp(\alpha_{y'} + \mathbf{x}_i \beta_{y'} + \mathbf{w}_{g,i} \mathbf{X}_g \delta_{y'} + \mathbf{w}_{g,i} \mathbf{p}_{gy'}^e \gamma_{y'} + v_{g,y'})}. \quad (2)$$

In Manski (1993)'s terms, γ_y is the endogenous effects, δ_y is the exogenous effects, and $v_{g,y}$ is the correlated effects. An individual's choice is affected by the expected behavior of peers (i.e. the endogenous effect) and their characteristics (i.e. the exogenous effect). The correlated effect may arise through endogenous group formation, common institutional or environmental factors which cause group members to behave similarly even in the absence of social effects.

Similarly to Lee et al. (2014), we assume agents face network interactions mediated by $\mathbf{W}_g$. Individuals form beliefs about the choice of their peers taking into account their characteristics

⁴With a bit abuse of notation we use g both as the social group index and the set of all individuals belonging to that group.

⁵In our application the strength of the tie is measured as a function of the geographical distance separating two households, and neighbors are individuals belonging to the same group, within a 50 meters distance from each other.

⁶The vector $\mathbf{p}_{gy}^e = (p_{gy,1}^e, \dots, p_{gy,n_g}^e)'$ is the belief any individual in group g has on j taking action y

⁷As is standard in the literature (Blume et al., 2010; Anderson et al., 1992) we impose a normalization on the dispersion in the random utility term equal to 1 and zero location parameter. This assumption amounts to homoskedasticity.

⁸Due to our rational expectations assumption $\mathbf{p}_{gy}^e$ should be equal to $\mathbf{p}_{gy} = (p_{gy,1}, \dots, p_{gy,n_g})'$.

and weighting these characteristics by the strength of their ties. This is especially relevant in smaller communities where we might expect individuals to personally interact with each other and therefore know the characteristics of specific peers. In contrast, Brock and Durlauf (2006) assume agents within a group possess homogeneous rational expectations due to group interactions: every individual is linked to everybody else, attaching equal weight to their influence (i.e. $w_{g,ij} = w \neq 0 \forall j \in g \setminus \{i\}$). Under the homogenous assumption on rational expectations, individuals observe the characteristics of the group as a whole. Given that every individual has the same information set, they all agree on their beliefs about the actions taken by everyone else.

The vector $\mathbf{p}_{gy}$ is the fixed point solution to the following expression

$$\mathbf{p}_{gy} \equiv \begin{pmatrix} p_{gy,1} \\ \vdots \\ p_{gy,n_g} \end{pmatrix} = \begin{pmatrix} \frac{\exp(\alpha_y + \mathbf{x}_1 \beta_y + \mathbf{w}_{g,1} \mathbf{X}_g \delta_y + \mathbf{w}_{g,1} \mathbf{p}_{gy} \gamma_y + \nu_{g,y})}{\sum_{y' \in Y} \exp(\alpha_{y'} + \mathbf{x}_1 \beta_{y'} + \mathbf{w}_{g,1} \mathbf{X}_g \delta_{y'} + \mathbf{w}_{g,1} \mathbf{p}_{gy'} \gamma_{y'} + \nu_{g,y'})} \\ \vdots \\ \frac{\exp(\alpha_y + \mathbf{x}_{n_g} \beta_y + \mathbf{w}_{g,n_g} \mathbf{X}_g \delta_y + \mathbf{w}_{g,n_g} \mathbf{p}_{gy} \gamma_y + \nu_{g,y})}{\sum_{y' \in Y} \exp(\alpha_{y'} + \mathbf{x}_{n_g} \beta_{y'} + \mathbf{w}_{g,n_g} \mathbf{X}_g \delta_{y'} + \mathbf{w}_{g,n_g} \mathbf{p}_{gy'} \gamma_{y'} + \nu_{g,y'})} \end{pmatrix}, \quad (3)$$

If we collect the L -left hand side vectors from expression (3), we get the $n_g \times L$ matrix $\mathbf{P}_g = (\mathbf{p}_{g0}, \dots, \mathbf{p}_{gL-1})$. Denoting the right hand side as $\Psi(\cdot)$ we get

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

$$\text{where } \theta = \left(\alpha_y, \beta_y, \delta_y, \gamma_y, (\nu_{g,y})_{g \in G} \right)_{y \in Y}$$

There are two main threats to the identification of the parameters θ . First, there is the standard problem of non-random sorting of individuals into groups. Individuals choose which group they would like to belong to (i.e. $g \in G$) and with whom they would like to interact (i.e. $\mathbf{w}_{g,i}$). The resulting correlation in unobservables among peers can lead to serious bias in the estimation of social interaction. This self-selection problem can be dealt with by operating under random assignment within *and* across groups based on observables along with group-fixed effects (e.g. Sacerdote (2001); Brock and Durlauf (2007); Lin (2010); Zimmerman (2003)). It can also be addressed by imposing exclusion restrictions on the structural model (e.g. Evans et al. (1992); Graham and Hahn (2005)) or by introducing variance restrictions on the error terms that are independent to the group size (Graham, 2008).

Second, the presence of correlated effects, created by common unobserved information shocks that hit the group as a whole, prevents the separate identification of the exogenous effects from the endogenous effects in the linear-in-means framework with symmetric influence (Manski, 1993; Brock and Durlauf, 2001; Blume et al., 2010; Moffitt, 2001).⁹ We show below that even in the presence of correlated effects, if there is enough variation in the adjacency matrix across rows, it is possible to separately identify the endogenous effects from the exogenous effects. This provides an extension to Brock and Durlauf (2006)'s symmetric influence case. As they do, we do not impose the rational expectation condition between the expected behaviors and the characteristics of a group, but rather we take these vectors as known regressors.

⁹When studying a linear-in-means model, the identification is further complicated by the simultaneity problem, also known as the reflection problem (Manski, 1993). However, the non-linear functional form given by the discrete choice model breaks this simultaneity problem (Brock and Durlauf, 2001). For instance, in the symmetric influence multinomial case without group unobservables, Blume et al. (2010)'s Theorem 13 provides sufficient conditions for the identification of θ up to a normalization on one of the alternatives.

Proposition 1. *Under $L > 2$ and the following assumptions*

- A.1 *Joint support of $(\mathbf{x}_i, \mathbf{w}_{g,i}\mathbf{X}_g)$ 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 *For each y , there is a group g such that conditional on $\mathbf{W}_g\mathbf{X}_g$, $\mathbf{x}_i$ is not contained in any proper linear subspace of $\mathbb{R}^K$,*
- A.4 *None of the elements of $\mathbf{x}_i$ contains bounded support,*
- A.5 *For each y , across different groups, $\mathbf{p}_{g,y}$ and $v_{g,y}$ are not constant,*
- A.6 *There is a group g where there is at least a pair $i, j \in g, i \neq j$, such that $\mathbf{w}_{g,i} \neq \mathbf{w}_{g,j}$.*

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

This result implies that after ruling out collinearity between regressors, imposing sufficient within-group variation on choices, characteristics and the weighting matrix in at least one group, the structural parameters are identified up to some normalization, even in the presence of correlated effects. More specifically, there should be at least a pair of individuals within a group that weighs differently the other group members. This 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. Notice that, given the non-linearities at the core of the discrete choice model (see Brock and Durlauf, 2006), the requirement on the structure of the network is weaker than for the continuous case (Bramoullé, 2013).¹⁰

Contrary to linear models of social interactions, our framework exhibits multiple equilibria as the structure is compatible with more than one aggregate outcome. Proposition 2 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 γ for which a unique equilibrium exists.

Proposition 2. Multiplicity. *In the multinomial choice model with asymmetric influence and network interactions given by (1) and (4), assuming $\gamma_y = \gamma, \delta_y = \delta, \alpha_y = \alpha$ for all $y \in Y$ and abstracting from the effect of the F_X on choice probabilities, if $|\gamma| < 4\left(1 - \frac{1}{L}\right)$ then there is a unique equilibrium.*

This result is a direct extension to Lee et al. (2014)’s condition for the binary case. It is worth highlighting that our identification result does not require a unique equilibrium given that we are assuming that $\mathbf{p}_{g,y}$ is observed by the econometrician.¹¹ We will come back to this discussion in the next section where we consider the estimation procedure.

3. Estimation Procedure and Simulation

Denoting $\mathbf{X}$ as the exogenous observables, $\mathbf{W}$ as the observed weighting matrix, $\mathbf{P}$ as the collection of $\mathbf{P}_{g,y} \forall g, y$ and v as a fixed effect at the group g level, the pseudo log-likelihood function is

¹⁰The details of the proofs can be found in the Appendix A.

¹¹For a review of the literature on the estimation of games with multiple equilibria see De Paula (2013).

$$L_N(Y | \mathbf{X}, \mathbf{W}, \mathbf{P}; \theta) = \frac{1}{N} \sum_{g \in G} \sum_{i \in g} \log \left[\frac{\sum_{y \in Y} \left(\exp(\alpha_y + \mathbf{x}_i \beta_y + \mathbf{w}_{g,i} \mathbf{X}_g \delta_y + \mathbf{w}_{g,i} \mathbf{p}_{gy} \gamma_y + \nu_{g,y}) \right) \mathbb{1}_{[y_i=y]}}{\sum_{y' \in Y} \exp(\alpha_{y'} + \mathbf{x}_i \beta_{y'} + \mathbf{w}_{g,i} \mathbf{X}_g \delta_{y'} + \mathbf{w}_{g,i} \mathbf{p}_{gy'} \gamma_{y'} + \nu_{g,y'})} \right]. \quad (5)$$

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

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

For computational reasons, various estimation procedures have been proposed in the literature (Aguirregabiria and Mira, 2007; Hotz and Miller, 1993; Pesendorfer and Schmidt-Dengler, 2008). Aguirregabiria and Mira (2007) suggested a Nested Pseudo Maximum Likelihood Estimation procedure (NPL algorithm). The first step is to find an estimator for $\mathbf{p}_{gy}$, denote it $(\hat{\mathbf{p}}_{gy}^0)_{g \in G, y \in Y}$ and fix $\hat{\mathbf{P}}^0$. The next step is to implement the Pseudo Maximum Likelihood (ML) maximization using a Newton–Raphson algorithm, eliminating the need to solve the fixed point problem because the estimated parameters in iteration- t are used to update beliefs using $\hat{\mathbf{P}}^t = \Psi(\hat{\mathbf{P}}^{t-1}, \mathbf{X}, \mathbf{W}; \hat{\theta}^t)$. This iterative process combines ML maximization with fixed-point updating until $\hat{\theta}^t$ is within a level of tolerance with respect to $\hat{\theta}^{t-1}$.¹²

Kasahara and Shimotsu (2012) propose a log–linear combination of $\Psi(\hat{\mathbf{P}}^{t-1}, \mathbf{X}, \mathbf{W}; \hat{\theta}^t)$ and $\hat{\mathbf{P}}^{t-1}$ given by

$$\Lambda^t = \left\{ \Psi(\hat{\mathbf{P}}^{t-1}, \mathbf{X}, \mathbf{W}; \hat{\theta}^t) \right\}^\phi \hat{\mathbf{P}}^{t-1^{1-\phi}} \quad (6)$$

with $\phi \approx 0$, such that the Λ^t –mapping has a local contraction property. This is known as the Relaxation Method (i.e. *NPL- Λ algorithm* in Kasahara and Shimotsu (2012)). It is shown that this estimation procedure converges to a consistent estimator even when the fixed point mapping does not have a local contraction property. Given our proposition 2, there could be no contraction property around our true parameters. We therefore follow the *NPL- Λ algorithm*.

Allowing for correlated effects at the group level $\nu_{g,y}$ induces an incidental parameters problem which leads to the inconsistency of maximum likelihood estimators (Neyman and Scott, 1948). This arises from the fact that 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 effects is not feasible as it introduces too many fixed effect parameters to estimate.¹³

¹²Bisin et al. (2011) advises implementing such recursive method for $T = 2$ iterations. Lee et al. (2014) substitute the fixed point updating step with $\hat{\mathbf{P}}^t$ being the solution to the fixed point iteration $\hat{\mathbf{S}}^t = \Psi(\hat{\mathbf{S}}^t, \mathbf{X}, \mathbf{W}; \hat{\theta}^t)$.

¹³To deal with this concern, Lee et al. (2014) propose to account for correlated effects by including both fixed effects at a broader level group and a random effect at a group level. Another possible strategy is to base the estimation on a conditional maximum likelihood 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). One needs to construct a likelihood function that conditions on a sufficient statistic for the incidental parameter. 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). Boucher et al. (2014) give the first empirical application of Lee (2007)’s results. 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.

Using Montecarlo simulations we investigate how acute the incidental parameters problem may be for the NPL- Λ algorithm. The details of the procedure along with the table of results are in the Appendix B. The results indicate that the larger the number of groups and the larger their size, the closer the estimates are to the true parameters and the smaller is the dispersion. This is reassuring given our empirical application where we have 186 different groups, with an average of 692.0 members per group and a mean of 20.33 neighbors per group member (see Appendix C). One important result is that estimates of the endogenous effect for a given alternative y is very precise when there is a large mass of individuals choosing that alternative, because $\hat{\mathbf{p}}_{gy}^t$ will be a consistent estimator of $\mathbf{p}_{gy}$ as $t \rightarrow \infty$.

4. Empirical Application

We apply our model to nineteenth century London where individuals choose their labor occupation and are influenced by the members of their community. In this section, we first provide a historical background to justify our definition of community based on ecclesiastical parish boundaries. After presenting our newly constructed data, we present the results. We finally perform several robustness checks.

4.1. Historical Background

During the Victorian era ecclesiastical parishes were 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 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).

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 (Pendrill, 1937). 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 (Firth, 1888).¹⁴ This Act established the Metropolitan Board of Works and empowered it to develop and implement schemes of London-wide significance. It also created 15 local Board of Works (BW), which were groupings of 55 of the smaller parishes, with statutory powers to manage and improve local facilities such as streets, paving, lighting, drainage and sewerage.¹⁵ 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 (also known as the

¹⁴Prior 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.

¹⁵The responsibility of the rearrangement of London’s local authorities fell into the hands of Cabinet Minister being “able to rearrange the boundaries of London unions at discretion” (The Economist, June 19, 1869)

Forster Act) relieved part of the education role which was previously under the control of the established church.

Despite losing importance in terms of civic responsibilities, its social importance remain intact: 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* or other spiritual benefits, right to have one's marriage solemnized, etc.

The institutional layout of Victorian London means that both ecclesiastical parish and BW 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 (Green, 1982; Johnson and Pooley, 1982; Green, 1991; Hoggart and Green, 1991). Given that parish membership was determined by domicile¹⁶, we use the ecclesiastical parish boundaries as proxies for social communities. In addition to the religion aspect of interaction, this definition of social group also captures the fact that social interactions were very local in nature in the late nineteenth century London. The distances over which most people travelled to work remained relatively short.¹⁷ We use the spatial distance between individuals as a proxy for the strength of their ties. Through their civil services, it can be argued that BW were instead relevant for Londoners' choice of residential location.

4.2. *Data and Descriptive Statistics*

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 (Schürer and Woollard, 2003) from the North Atlantic Population Project (NAPP, 2015). The unit of observation is at the individual level. The census contains the full address of individuals (house number or name, name of street, avenue or road, civil parish and county of residence). In addition to geographic variables, the census also provides a wider range of sociodemographic information: age, gender, place of birth, marital status, number of children, number of servants and family structure as well as information on self-reported occupation. There are over 400 occupations. Following Woollard (1998), we have aggregated occupations into three categories: professional, commercial, and industrial.¹⁸ Professional workers include lawyers, doctors and teachers, and

¹⁶Alternatively, parish membership was established through particular groups such as ethnic parishes, college parishes, etc.

¹⁷The mean journey to work for those employed in London was only around 5 km 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). 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 a large number of people (Dyos, 1953; Green, 1988; Lawton, 1959; Warnes, 1972).

¹⁸Occupations are broadly defined so that they are not perfect substitutes. There are also domestic occupations and unemployed category, each representing less than 6% of the male population. We do not include this population in the

commercial workers include merchants, dealers of money and drivers. Industrial workers are further divided into artisan, builder, food/agriculture and services.

Using historical maps, we geo-reference all the streets of 1881 London. We first extend the digitalized map of London dating back to John Rocque's 1746, which was provided by the Archaeology section of the Museum of London. We then manually add points for each street using the 1882 First Ordnance Survey Map of London. We have a total of 5,998 referenced streets. Finally, we overlay the ecclesiastical parish and BW boundaries provided by the UK Data Service. There are 299 ecclesiastical parishes in Central London and 38 BW.

We geographically locate individuals in the census on the historical map using their place of residence (address, parish and county) and the geo-referenced street points along with the ecclesiastical and BW boundaries. Our final merged dataset comprises 1,137,876 individuals located down to the street level. This amounts to a match for 70% of the entire population in London in 1881. Our sample focuses on native men of working age (between the ages of 15 and 60) who are household heads. We eliminate foreign-born individuals and individuals who live in the place where they work such as prisons, workhouses or other public institutions. This is to ensure that individuals are aware of the institutional layout of London and are integrated in a social group. We also 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 neighbor living on the same street. These restrictions correspond to the simulation results in the previous section and agree with standard procedure in the literature to avoid small sample issues that may affect the estimates. We end up with a total of 186 ecclesiastical parishes within 32 BW. Generally, we have large variation across and within BW. In figure C.1 of Appendix C, we show the map of London ecclesiastical parishes included in our analysis (i.e. 186 parishes) and the distribution of the number of ecclesiastical parishes per BW (mean 5.81, sd dev 4.3), of the members per parish (mean 692.0 sd dev 505.8) and of the average number of neighbors living in the same street per parish (mean 20.34, sd dev 24.1).

Table 1 reports the descriptive statistics by occupational categories. 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 C.2 in Appendix C map these various characteristics. The eastern part of London was predominantly inhabited by younger men. Larger families predominantly lived in the south. Wealth, captured by the number of servants, was mainly found in the west. The majority of individuals living in the west were born elsewhere. The employment structure of London was diverse with industrial occupation dominating the labor market. Based on our occupational classification, 6.5% of the sample worked in a professional occupation, 21.41% worked in a commercial occupation and finally 75.97% held industrial jobs (with a minimum of 13.5 % for industrial occupations related to services, and a maximum of 23.7 % on artisans within industrial occupations). We also show the geographic pattern by occupations in figures C.3 in Appendix C. As already documented by Topa and Zenou (2014) in other setting, there was also a spatial clustering of occupations within London. Employment appeared to be predominant in the central areas of London. Professional trades accounted for a large proportion of West London. In contrast, commercial occupations were concentrated in East London. Finally, industrial workers

analysis. This responds, on one hand, to the fact that people working in domestic occupations generally live where they work, invalidating our measure of strength of ties. On the other, because we are studying occupational choices rather than decisions to participate in the labor force.

appeared to be more spread out.

Table 1: Summary statistics per Occupation, London 1881

	All	Professional	Commercial	Industrial			
				Artisan	Builder	Food/Agri	Services
Age	38.221 (10.40)	38.230 (10.16)	37.529 (10.31)	38.817 (10.64)	38.682 (10.35)	38.074 (10.29)	37.872 (10.37)
Married	0.959 (0.199)	0.910 (0.287)	0.961 (0.194)	0.964 (0.187)	0.971 (0.168)	0.954 (0.210)	0.962 (0.192)
# children	2.169 (2.050)	1.913 (2.011)	2.040 (1.979)	2.324 (2.112)	2.259 (2.066)	2.055 (2.008)	2.261 (2.072)
# servants	0.175 (0.813)	0.662 (1.699)	0.106 (0.711)	0.116 (0.652)	0.061 (0.384)	0.339 (1.026)	0.084 (0.455)
Residents							
Parish birth	0.136 (0.343)	0.059 (0.236)	0.118 (0.322)	0.155 (0.362)	0.164 (0.370)	0.130 (0.336)	0.143 (0.350)
County birth	0.474 (0.499)	0.323 (0.468)	0.460 (0.498)	0.504 (0.500)	0.497 (0.500)	0.462 (0.499)	0.503 (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 a parish which has a minimum of 30 residents within a BW which has at least two ecclesiastical parishes. Working categories by Professional: Civil service, Clerical, Legal professions, Medical professions, Education, Liberal arts, Scientists and Sports; by Commercial: Sales, Transport and communication; Industrial Artisans; Industrial Builders; by Industrial Food/Agro: Agriculture and Food Dealer; by Industrial Services: Services, Sales, Media and Technical.

4.3. Results

In table 2 we present our estimation results, following the decision model expressed in equations (2)-(3), where each Londoner, taking social group as given, chooses an occupation among professional, commercial or the different types of industrial jobs (i.e. artisan, builder, food and agriculture, or services). We define $\mathbf{W}_g$ as 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. We control for individual characteristics x_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. Network-level covariates $\mathbf{w}_{g,i}\mathbf{X}_g$ include the same characteristics aggregated at the social group level and weighted by the strength of ties between individuals of the group.

Table 2: Estimation of endogenous effects, γ_y , and Average Marginal Effects times standard deviation of estimated beliefs

Occupation (y)	NPL- Λ		AME [†]
	(1)	(2)	(3)
Professional	2.341*** (0.672)	1.843** (0.753)	0.72% [0.001]**
Commercial	-11.146*** (2.761)	-9.668*** (2.749)	-9.93% [0.013]***
Industrial Artisan	2.343** (0.983)	1.7* (1.002)	1.96% [0.009]*
Industrial Builder	3.526*** (0.994)	3.279*** (1.113)	2.71% [0.006]***
Industrial Food/Agro	2.213** (0.864)	2.773*** (0.698)	2.26% [0.003]***
Industrial Services	1.921 (2.864)	0.313 (3.505)	0.14% [0.009]
log-like	-214,390	-212,190	
AIC:	429,210	426,350	
obs	128,709		
Individual characteristics	yes	yes	yes
Group characteristics	yes	yes	yes
BW fixed effects	yes	no	no
Parish fixed effects $\nu_{g,y}$	no	yes	yes

Notes: Robust Standard errors clustered at parish level in parentheses. Krinsky and Robb (1986)'s Standard errors after 1,000 replications in square brackets. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. Table reports endogenous effects γ_y , $\forall y \in Y$ from equations (2)-(3) model, where occupation choice set Y is professional, commercial, industrial artisan, industrial builder, industrial in food and agriculture, or industrial in services. Individual characteristics x_i include age, sex, 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- Λ uses $\phi = 0.1$. [†] AME stands for Average Marginal Effect times the standard deviation of the weighted estimated beliefs.

We present the NPL- Λ structural estimation with $\phi = 0.1$.¹⁹ As noted previously, this algorithm converges to the true parameters even when the fixed point constraint (4) does not have local contraction properties in a neighborhood of the true parameters. In column 1, we control for fixed effects at the BW level. Given the role of public good services provider of BW, it can be argued that individuals' location decision was based on BW. Adding BW fixed effect therefore deals in part with the issue of self-selection into groups.²⁰

¹⁹We also undertook the estimation with $\phi = 0.8$ and results do not change significantly.

²⁰In the absence of random peer groups (Hoxby, 2000; Sacerdote, 2001) most studies incorporate group-specific fixed

To gain insight into the correlated effects at the group level (i.e. ecclesiastical parish), we add parish fixed effect in column 2. These correlated effects can take the form of local industries or an inspiring priest encouraging his parishioners to work or share information. Failing to take the correlated effects into account can lead to an upward bias in the estimation. In column 3, we follow Lee et al. (2014) and present the Average Marginal Effects times the standard deviation of the weighted estimated beliefs (AME) associated to the specification in column 2. We provide more details on the calculation of marginal effects in Appendix D.²¹

We find that social groups play a significant and positive role for individuals in professional and industrial occupations. A one standard deviation change in the expected ratio of peers in professional occupations leads to a 0.72% increase in the likelihood of being employed in a professional occupation. Similarly, a one standard deviation change in the expected ratio of peers in industrial occupations leads to an increase in the likelihood of working in an industrial occupation ranging from 0.14% to 2.71%. Professional and industrial occupations are mainly demand-side driven and information about job opening should therefore be easily transmitted. In contrast, commercial occupations tend to be more competitive. Individuals may want to keep the information they possess private. Thus, we find that a one standard deviation increase in the expected ratio of peers in commercial occupations reduces the chance of following a similar occupation by 9.93%. Notice that such magnitude is somewhat lower than those found in contemporary studies. Topa (2001) finds that a one standard deviation of peers' employment leads to an increase in the likelihood of being employed between 0.6% and 1.3%, while Bayer et al. (2008)'s estimates lie between 0.8% and 3.6%.²²

4.3.1. Robustness Checks

Definitions of $\mathbf{W}_g$ and Placebo Coordinates

In the same vein as Lin (2010), we first explore how sensitive our results are to different specifications of the adjacency matrix. In the first two columns of table 3, we define the adjacency

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 decide which school to send their children to but cannot decide which class they will be part of. Thanks to the unique administration layout of Victorian London, it is plausible that individuals choose their residential location based on the public services provided at the BW level. Therefore adding BW fixed effects partially controls for self-selection into a group.

²¹The estimated standard errors for the AME follow the Krinsky and Robb (1986) method which draws multiple vectors of θ_s , $s = 1, \dots, S$ coefficients, from the multivariate normal distribution that has a mean vector equal to the estimated coefficient vector $\hat{\theta}$ and the same estimated variance-covariance, from the *NPL- Λ algorithm*. Each new vector of coefficients, θ_s , is used to compute a new value of $AME_{y,s} = \hat{\mathbb{E}}_i \left[\frac{d\hat{p}_{y,i}}{d\mathbf{w}_i\hat{\mathbf{p}}_y} \hat{d}_i(\mathbf{w}_i\hat{\mathbf{p}}_y); \theta_s \right]$. The standard deviation of the resulting sample of draws of $AME_{y,s}$ across all values of θ_s provides an estimate of the standard error for the AME. We use $S = 1,000$ replications (see Dowd et al. (2014) for a discussion on how this method compares to bootstrapping and delta methods).

²²In Appendix E, we present the average marginal effects of the exogenous effects. As explained in Appendix D, an individual's characteristic affects the probability of choosing an occupation y via a "direct effect". But also, individual's covariates affect the expected behavior of her parish members. Therefore, there are also "indirect effects": one coming from the change in her peers' probability of following occupation y , and a second from the change in the expected behavior of her peers following any other occupation $y^0 \neq y$. Among these exogenous variables, we find that individual's characteristics direct effects are significantly more important than the indirect ones. Age, number of servants and being married have a negative direct effect on individuals for any occupation category, while number of children and living in the parish of birth has a positive effect.

matrix as $w_{g,ij}^\delta = 1/|nei_i^\delta| \forall j \in g$ such that $\|i - j\| \leq \delta$ meters, where index i is used as a label for an individual as well as her coordinates. We use $\delta = \{0, 100\}$. For a distance of 0 meter, where we only consider individuals from the same street, results are very similar to our main specification.²³ Once we allow for a 100 meters radius within the same parish we see that the endogenous effects are smaller in magnitude suggesting that closer neighbors have larger effects on individual's occupational choice.

Table 3: Estimation of endogenous effects, γ_y , with various adjacency matrices and different samples

Occupation (y)	NPL- Λ					
	$w_p^{\delta=0}$	$w_p^{\delta=100}$	Placebo coords within		Non	Ages
	(1)	(2)	BW	London	movers	15-30
Professional	1.721** (0.705)	1.708** (0.844)	3.945 (2.533)	3.818 (3.854)	1.389 (1.587)	0.292 (1.07)
Commercial	-7.046*** (2.389)	-5.327 (3.292)	-0.763 (3.297)	0.127 (4.745)	-10.065*** (3.435)	-1.862 (2.343)
Industrial Artisan	1.492 (1.115)	2.86*** (1.081)	2.313 (1.93)	-0.301 (5.428)	-0.842 (2.106)	1.316 (2.185)
Industrial Builder	3.318*** (1.026)	2.475* (1.401)	3.95* (2.361)	3.324 (2.901)	3.246*** (2.325)	3.128 (1.18)
Industrial Food/Agro	3.019*** (0.639)	3.057*** (1.027)	3.689* (1.895)	3.767 (2.858)	3.02*** (1.245)	2.783** (0.957)
Industrial Services	0.475 (3.001)	-0.768 (4.51)	2.768 (1.951)	0.961 (8.08)	1.176 (4.273)	0.213 (3.293)
Log-Likelihood:	-212,190	-212,210	-212,840	-217,360	-98,104	-59,744
AIC:	426,350	426,380	427,660	436,690	1,839	121,280
obs		128,709			60,796	36,339
Individual characteristics	yes	yes	yes	yes	yes	yes
Group characteristics	yes	yes	yes	yes	yes	yes
Parish fixed effects $v_{g,y}$	yes	yes	yes	yes	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 γ_y , $\forall y \in Y$ from equations (2)-(3) model, where occupation choice set Y is professional, commercial, industrial artisan, industrial builder, industrial in food and agriculture, or industrial in services. Column 1 (2) defines neighbors as two individuals living in the same street (within 100 meters). Column 3 (4) randomly allocates individuals within BW (across London). Column 5 restricts sample to individuals that live where they were born. Column 6 to individuals between 15-30 years old. Individual characteristics x_i include age, sex, 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- Λ uses $\phi = 0.1$.

Another concern is the potential measurement error in the location of each address on the historical map. In the absence of postal codes, we located individuals based on their street address and parish of residence. We cannot locate exact house numbers. The point referencing a street address 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. In an attempt to examine how this aggregation could have influenced the statistical

²³Notice that our parameters are identified provided that, within a parish, streets have different numbers of residents, thus guaranteeing that the rows of the adjacency matrix vary for two residents living in different streets.

significance of our results, we implement two placebo tests. In the first, we randomly allocate individuals to different streets *within* the same BW. The estimates from this placebo test are included in column 3. The endogenous effects keep their signs and become insignificant for nearly all occupations. For industrial builders and industrial food and agriculture, the significance level is lowered. This reveals the importance of the weighting matrix defined at the parish level. The endogenous effects appear to be driven by social group interactions as opposed to possible neighborhood unobservables. We also randomly allocate individuals to different points across London. Results of this second placebo test are presented in column 4. As expected, results become insignificant and change significantly in sign and magnitude. Thus we are reassured that our results rely on correctly identifying social groups.

Lastly, we restrict the sample to individuals who are living within the same county where they were born (column 5) and to younger cohort aged between 15 and 30 years old (column 6). In both specifications, the sign of the endogenous effects remain the same as in the baseline specification. When looking at the sample of individuals who lived where they were born, the endogenous effect is no longer significant for professional occupations. This might be explained by a possible migration response to the lack of job opportunities in this sector. The results on younger cohorts show that the endogenous effects have the same signs as in the baseline but are no longer significant for the majority of occupational categories. This suggests that peers from the same age group is not necessarily the appropriate reference group for younger cohorts.

Multiplicity

As stated in proposition 2, multiple equilibria could arise whenever endogenous effects are larger than 3.33 in absolute value. Column 2 in table 2 shows that commercial occupations have a larger endogenous effect than this threshold. We investigate whether there are multiple equilibria given our estimated parameters. For each parish g we find the roots of the large $(n_g(L - 1))$ -system of non-linear equations described by

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

To do so we follow spectral methods to solve for possible multiple roots.²⁴ However, none of the 186 parishes included in our study have more than one equilibrium.²⁵

5. Conclusion

This paper presents the identification and estimation of 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 adjacency matrix. We provide a sufficient condition for a unique equilibrium. We

²⁴We implemented the spectral method in *R-package* following Varadhan and Gilbert (2009). Results are available upon request.

²⁵After using a sufficiently large number of different starting values per each parish i.e. 1005.

use a Relaxation Method proposed by Kasahara and Shimotsu (2012) to provide consistent and asymptotically efficient estimates of the structural parameters.

As an empirical application, we investigate how social groups affected occupational choices in Victorian London. We construct a new dataset which geographically locates London residents enumerated in the full census of 1881. We argue that ecclesiastical parishes were a defining feature of social networks and use ecclesiastical parish boundaries as proxies for social groups. Our results indicate that social interactions within parishes play a role in determining labor market outcomes. Accounting for correlated effects at the group level, we find that an increase in the share of peers employed in an industrial or professional occupations increases an individual's probability of being employed at that same occupation, while for commercial occupations the reverse is true.

While our application relies 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.

The strong endogenous effects we find suggest that any program targeting employment rates in particular sectors has the potential to have spillovers effects through social network. This may have implications both for social mobility and the industrialization process. For instance, inter-generational mobility in terms of occupations may remain low because workers seek to use their inherited social connections to find jobs more easily as documented by Borjas (1994) and Munshi and Wilson (2008). This in turn has consequences for the spatial patterns of income inequalities within a city.

Appendix A. Proofs

Proof of Proposition 1

Proof. When normalizing common parameters for $y = 0$ to be 0 (i.e. $\alpha_0 = 0, \beta_0 = \delta_0 = \mathbf{0}$ and $\nu_{g,0} = 0$ for every g), we know that,

$$\log\left(\frac{p_{gy,i}}{p_{g0,i}}\right) = \alpha_y + \mathbf{x}_i\beta_y + \mathbf{w}_{g,i}\mathbf{X}_g\delta_y + \mathbf{w}_{g,i}(\mathbf{p}_{gy}\gamma_y - \mathbf{p}_{g0}\gamma_0) + \nu_{g,y}.$$

By contradiction, we assume there is another set of observationally equivalent structural parameters $\bar{\theta}$, then it must be the case that

$$\mathbf{x}_i(\beta_y - \bar{\beta}_y) + \mathbf{w}_{g,i}\mathbf{X}_g(\delta_y - \bar{\delta}_y) + \mathbf{w}_{g,i}\mathbf{p}_{gy}(\gamma_y - \bar{\gamma}_y) + \mathbf{w}_{g,i}\mathbf{p}_{g0}(\bar{\gamma}_0 - \gamma_0) = \bar{\alpha}_y - \alpha_y + \bar{\nu}_{g,y} - \nu_{g,y}.$$

Notice that for a given g we have, for every given y , that the right hand side remains constant while there is variation, due to A.1 and A.6, on the left hand side. Then, for the equality to hold, it must be the case that $\alpha_y + \nu_{g,y} = \bar{\alpha}_y + \bar{\nu}_{g,y}$ where we can apply the second part of A.5 and get $\alpha_y = \bar{\alpha}_y$ and $\nu_{g,y} = \bar{\nu}_{g,y}$. Then

$$\mathbf{x}_i(\beta_y - \bar{\beta}_y) + \mathbf{w}_{g,i}\mathbf{X}_g(\delta_y - \bar{\delta}_y) + \mathbf{w}_{g,i}\mathbf{p}_{gy}(\gamma_y - \bar{\gamma}_y) = \mathbf{w}_{g,i}\mathbf{p}_{g0}(\gamma_0 - \bar{\gamma}_0).$$

For a given duple (g, i) by the first part of A.5, and considering $L > 2$, for the equality to hold for every y it must be the case that $\gamma_0 = \bar{\gamma}_0$. Which leaves us with

$$\mathbf{x}_i(\beta_y - \bar{\beta}_y) = \mathbf{w}_{g,i}(\mathbf{X}_g(\bar{\delta}_y - \delta_y) + \mathbf{p}_{gy}(\bar{\gamma}_y - \gamma_y)).$$

Notice that from A.2 and A.3, if we fix a group g for every y , the previous equality can hold if and only if $\beta_y = \bar{\beta}_y$. Then it must be the case that

$$\mathbf{w}_{g,i}\mathbf{X}_g(\delta_y - \bar{\delta}_y) = \mathbf{w}_{g,i}\mathbf{p}_{gy}(\bar{\gamma}_y - \gamma_y).$$

We know that A.4 imply that the LHS support is unbounded, but given (3) we know that each element $p_{gy,i} \in [0, 1]$, so it must be the case that $\delta_y = \bar{\delta}_y$. As we know by A.6 that for each y , $\mathbf{w}_{g,i}\mathbf{p}_{gy}$ varies across individuals so it must also be true that $\bar{\gamma}_y = \gamma_y$. $\square$

Proof of Proposition 2

Proof. Let us leave group g and choice-group unobservables conditioning implicit. We know that $p_{y,i} = \frac{\exp(\alpha_y + \mathbf{x}_i\beta_y + \mathbf{w}_i\mathbf{X}\delta_y + \mathbf{w}_i\mathbf{p}_y\gamma)}{\sum_{y' \in Y} \exp(\alpha_{y'} + \mathbf{x}_i\beta_{y'} + \mathbf{w}_i\mathbf{X}\delta_{y'} + \mathbf{w}_i\mathbf{p}_{y'}\gamma)}$. Therefore

$$m_{y,i} \equiv p_{y,i} - p_{0,i} = \frac{[\exp(g_{y,i} + \mathbf{w}_i\mathbf{m}_y\gamma) - 1]}{1 + \sum_{y' \neq 0} \exp(g_{y',i} + \mathbf{w}_i\mathbf{m}_{y'}\gamma)} \equiv \psi_{y,i}((\mathbf{m}_{y'})_{y' \in Y}, \gamma, (\mathbf{g}_{y'})_{y' \in Y}),$$

where $g_{y,i} \equiv \alpha_y - \alpha_0 + \mathbf{x}_i(\beta_y - \beta_0) + \mathbf{w}_i\mathbf{X}(\delta_y - \delta_0)$, $\mathbf{g}_y = (g_{y,1}, \dots, g_{y,n})'$, $\mathbf{m}_y = (m_{y,1}, \dots, m_{y,n})'$ and denote $\boldsymbol{\psi}_y = (\psi_{y,1}, \dots, \psi_{y,n})'$

We know that the $n \times (L-1)$ matrix $\mathbf{M} \equiv (\mathbf{m}_1, \dots, \mathbf{m}_{L-1}) = (\boldsymbol{\psi}_1, \dots, \boldsymbol{\psi}_{L-1}) \equiv \boldsymbol{\Psi}$.

If we assume $\mathbf{g}_y = 0 \forall y \in Y$ (i.e. we ignore the effect of X on choices and $\alpha_y = \alpha$ for all $y \in Y$), we get that $\mathbf{m}_y = 0 \forall y$ is an equilibrium. To see this, notice that $\psi_{y,i}((\mathbf{0})_{y' \in Y}, \gamma, (\mathbf{0})_{y' \in Y}) = \frac{\exp(\mathbf{w}_i\mathbf{0}\gamma) - 1}{1 + \sum_{y' \neq 0} \exp(\mathbf{w}_i\mathbf{0}\gamma)} = 0$ for every i, y .

Given this, consider the case for $\hat{\mathbf{M}} = (\mathbf{m}_1, \mathbf{0}, \dots, \mathbf{0})$ then $\mathbf{m}_1 = \boldsymbol{\psi}_1(\hat{\mathbf{M}}, \gamma, (\mathbf{0})_{y' \in Y}) = \frac{\exp(\mathbf{W}\mathbf{m}_1\gamma) - 1}{L-1 + \exp(\mathbf{W}\mathbf{m}_1\gamma)} \equiv \boldsymbol{\psi}_w(\mathbf{m}_1)$. 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 $\max_i \sum_j |a_{ij}|$. By the contraction mapping theorem we know that if there is a $r \in \mathbb{R}$ such that $0 \leq r < 1$ and $\|\boldsymbol{\psi}_w(\mathbf{m}_1) - \boldsymbol{\psi}_w(\mathbf{q}_1)\|_\infty \leq r\|\mathbf{m}_1 - \mathbf{q}_1\|_\infty$ then mapping $\boldsymbol{\psi}_w(\mathbf{m}_1)$ has a unique fixed point. By the Mean Value Theorem we also know 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 γ such that $0 \leq \|\nabla \boldsymbol{\psi}_w(\mathbf{m}'_1)\|_\infty < 1$. Notice that $\frac{\partial \boldsymbol{\psi}_w(\mathbf{m}'_1)}{\partial m_{1,i}} = 0$ given $w_{ii} = 0$ and $\frac{\partial \boldsymbol{\psi}_w(\mathbf{m}'_1)}{\partial m_{1,j}} = \frac{\exp(\mathbf{w}_j\mathbf{m}'_1\gamma)w_{ij}L\gamma}{L-1 + \exp(\mathbf{w}_j\mathbf{m}'_1\gamma)}$ for every $j \neq i$.

Therefore, $\|\nabla\psi_w(\mathbf{m}'_1)\|_\infty = \max_i \frac{|\gamma|L \exp(\mathbf{w}_i \mathbf{m}_1 \gamma)}{[L-1+\exp(\mathbf{w}_i \mathbf{m}_1 \gamma)]^2} \sum_{j \neq i} |w_{ij}| \leq |\gamma| \frac{L}{4(L-1)}$ which provides the result. The last inequality follows when one finds the maximum of the expression taking $\mathbf{w}_i \mathbf{m}_1 \gamma$ as the input variable. Notice the condition would be the same if we instead use the metric space $(\psi_w(\mathbf{m}_1), \|\cdot\|_1)$ where $\|A\|_1$ is the maximum absolute column sum norm of a matrix A given by $\max_j \sum_i |a_{ij}|$.

□

Appendix B. Montecarlo Simulation

To determine how acute the incidental parameters problem may be, we use a Montecarlo simulation. We simulated data with different number of groups $\#G \in \{5, 10, 50, 100\}$ with different group sizes $n_g \in \{50, 100, 300\}$. A world is a duple $(\#G, n_g)$. For each group g we simulate points on the space $[0, U_{[1,2]}]^2$ and define $w_{g,i}$ as all those individuals that are within distance r_g to i where r_g is the minimum distance such that every $i \in g$ has at least 10 neighbors. Individuals face the utility function given by equation (1), where we assume $K = 1$ and have to choose among five alternatives ($L = 5$).

For each reported estimation we generate 100 “worlds”. We assume that the Data Generation Process (DGP) is identical across different worlds. Even though the DGP is exactly the same for a given world, w_g and X_g are random so we have different choices across groups. Each simulation provides us with observational data on choices $\mathbf{Y}$, characteristics $\mathbf{X}$ and network $\mathbf{W}$ for every group $g \in G$. To simulate optimal choices with consistent asymmetric influence we solve, iteratively, for the belief fixed point equation (3) until no individual’s beliefs change. This guarantees only stable equilibria to emerge in the observational data. We report the estimation by the NPL- Λ Relaxation Method, described in section 3.

We assume the true coefficients are given by $\gamma^0 = (3.3, 2.5, 2, 3.2, 3.6)'$ (the endogenous effect), $\alpha^0 = (-1.5, -1.3, -2.4, -1.7, -2.1)'$ (alternative-specific coefficients), $\beta^0 = (1.5, 1.4, 2.1, 0.9, 1.1)'$ (individual characteristics) and $\delta^0 = (2.2, 2.4, 2.7, 2.3, 2.6)'$ (the exogenous effect). Regarding the correlated effect, we assume $\gamma_{g,y} \sim \sigma^0 N(0, 1)$ where $\sigma^0 = (0.13, 0.08, 0.18, 0.05, 0.1)$.²⁶

Table B.4 reports the median of the estimated coefficient of the endogenous effect and below, in parenthesis, the 25th and 75th percentile of the estimated coefficient for 100 simulations. The last row of the table shows the average percentage of individuals choosing each alternative. The results indicate that the larger the number of groups (i.e. $\#G$) and the larger their size (i.e. n_g), the closer the estimates are to the true parameters and the smaller is the dispersion. This is reassuring given our empirical application where we have 186 different groups, with an average of 692.0 residents and a mean of 20.34 neighbors per group member. One important result is that estimates of the endogenous effect for a given alternative is very precise when there is a large mass of individuals choosing that alternative.²⁷

²⁶The coefficients were chosen so as to have some alternatives being chosen by less than 10% of the population. The results remain valid for different sets of parameters.

²⁷Even though the consistency properties of such estimator are not analysed here, the result is reminiscent of the importance of the 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).

Table B-4: Montecarlo simulations and estimation of endogenous effect γ from simulated world with G groups, each of size n_g .

	(1)			(2)			(3)			(4)			(5)		
	$\gamma_0 = 3.3$ med			$\gamma_1 = 2.5$ med			$\gamma_2 = 2$ med			$\gamma_3 = 3.2$ med			$\gamma_4 = 3.6$ med		
n_g	50	100	300	50	100	300	50	100	300	50	100	300	50	100	300
#G	NPL-Λ, w observed														
5	3.01 (-3.7,6.1)	3.24 (-0.8,4.6)	2.89 (0.6,3.6)	2.73 (-2.3,4.4)	1.86 (-2.3,3.4)	2.17 (-0.3,3.1)	4.01 (-1.5,9)	2.37 (-0.4,3.9)	2.33 (0.6,3.4)	5.54 (-9.7,11.1)	4.64 (-6.5,7.9)	3.44 (-1.7,6.4)	7.81 (-0.7,12.5)	6.03 (-6.2,9.1)	3.43 (-8.1,8)
10	2.6 (-2.3,4.9)	3 (-1.4,1)	3.2 (2.5,3.6)	2.5 (-0.8,4.2)	2.6 (-0.5,3.2)	2.4 (1.3,3)	2.2 (-2.4,4.3)	2.5 (1.3,3.8)	2 (1.2,8)	2.7 (-5.4,8.7)	3.3 (-3.9,6)	3.8 (-0.2,5.2)	4.2 (-6.4,9.5)	5.1 (-3.6,8.3)	4.1 (-5.4,7.4)
50	2.4 (-0.7,3.8)	3.2 (2.6,3.6)	3.2 (3.3,4)	1.8 (-0.4,3)	2.2 (1.2,2.7)	2.3 (1.9,2.6)	2.4 (0.6,3.6)	2.1 (1.2,2.8)	2.1 (1.8,2.4)	2.4 (-3.3,5.1)	3 (0.4,6)	3.2 (2.4,2)	1.7 (-5.3,6)	3 (-4.9,5.9)	2.7 (-0.5,5.8)
100	3 (1.1,3.7)	3.1 (2.7,3.4)	3.2 (3.1,3.4)	1.8 (0.1,2.8)	2.3 (1.6,2.8)	2.5 (2.3,2.7)	1.8 (0.6,2.8)	1.9 (1.4,2.5)	2 (1.6,2.2)	2.5 (-1.2,4.9)	2.6 (1.5,3.8)	3.2 (1.8,3.8)	2.6 (-4.7,6.2)	3 (-1.5,5.1)	3.6 (0.8,5.4)
% choices	0.27			0.30			0.27			0.09			0.08		

Notes: *med*: median. The first number in parenthesis is the 25th percentile and the second the 75th percentile of the estimated coefficients. Table shows median, after 100 simulations, of the estimated coefficient of the endogenous effect γ_i from model expressed in equation (2)-(3). Every simulated data comes a duple of number of groups ($\#G$) and group size (n_g) and individual's randomly allocated on the Euclidean space $[0, U_{[1,2]}]^2$. Adjacency matrix $w_{g,i}$ is such that every $i \in g$ has at least 10 neighbors.

Appendix C. Further descriptives

(a) Parishes included in the study (gray excluded, white included)

(b) Number of parishes ($\#G$) per BW

(c) Size of parishes (n_g) per parish

(d) Mean of neighbors (nei_g) per parish

Figure C.1: Descriptive statistics on $\#G, n_g, nei_g$, London 1881

(a) Average age per parish

(b) Percentage of married individual per parish

(c) Average number of children per parish

(d) Number of servants

(e) Percentage of residents in parish of birth

Figure C.2: Mean of individual characteristics per ecclesiastical parish by quantile of the city distribution (red is upper quantile, blue is lower quantile)

(a) Professional Occupation

(b) Commercial Occupation

(c) Industrial Artisan Occupation

(d) Industrial Builder Occupation

(e) Industrial Food/Agro Occupation

(f) Industrial Services Occupation

Figure C.3: Percentage of individuals by occupation category at the street level by quantile of the city distribution (red is upper quantile, blue is lower quantile)

Appendix D. Marginal effects

Starting from equations (2) and (3) if we consider the change in the endogenous effect we have

$$\frac{dp_{y,i}}{d\mathbf{w}_i\mathbf{p}_{y^0}} = \begin{cases} \gamma_y p_{y,i}(1 - p_{y,i}) & \text{if } y^0 = y, \\ -\gamma_{y^0} p_{y^0,i} p_{y,i} & \text{if } y^0 \neq y. \end{cases}$$

However, for the individual characteristics we know a change in any of the covariates will have a direct effect on the probability of choosing occupation y (equation (2)) but also equilibrium, indirect, effects through equation (3). That is, an individual's characteristic affects the probability of choosing an occupation y via a "direct effect" (p.e.g. being older might reduce the probability of being in commercial occupation y). But also, an individual's covariate affects the expected behavior of her parish members. Therefore covariate changes would have also two "indirect effects": one coming from the change in her peers' probability of following occupation y , and a second from the change in the expected behavior of her peers following any other occupation $y^0 \neq y$. Therefore,

$$\frac{dp_{y,i}}{dx_{k,i}} = \underbrace{\frac{\partial p_{y,i}}{\partial x_{k,i}}}_{\text{direct effect}} + \underbrace{\frac{\partial p_{y,i}}{\partial \mathbf{w}_i\mathbf{p}_y} \frac{\partial \mathbf{w}_i\mathbf{p}_y}{\partial x_{k,i}}}_{\text{indirect } y: \text{ effect on beliefs of } j \in \text{nei}_i \text{ taking } y} + \underbrace{\sum_{y^0 \neq y} \frac{\partial p_{y,i}}{\partial \mathbf{w}_i\mathbf{p}_{y^0}} \frac{\partial \mathbf{w}_i\mathbf{p}_{y^0}}{\partial x_{k,i}}}_{\text{indirect } y^0 \neq y: \text{ effect on beliefs of } j \in \text{nei}_i \text{ taking any other } y^0 \neq y}$$

Doing the calculation we get

$$\begin{aligned} \frac{dp_{y,i}}{dx_{k,i}} = & \beta_{k,y} p_{y,i} \left(1 - \sum_{y' \in \Omega} \frac{\beta_{k,y'}}{\beta_{k,y}} p_{y',i} \right) + \\ & \gamma_y p_{y,i} (1 - p_{y,i}) \left[\sum_{j \in \text{nei}_i} w_{ij} \left(w_{ji} \delta_{k,y} p_{y,j} \left(1 - \sum_{y' \in \Omega} \frac{\delta_{k,y'}}{\delta_{k,y}} p_{y',j} \right) \right) \right] - \\ & \sum_{y^0 \neq y} \gamma_{y^0} p_{y^0,i} (p_{y^0,i}) \left[\sum_{j \in \text{nei}_i} w_{ij} \left(w_{ji} \delta_{k,y^0} p_{y^0,j} \left(1 - \sum_{y' \in \Omega} \frac{\delta_{k,y'}}{\delta_{k,y^0}} p_{y',j} \right) \right) \right]. \end{aligned}$$

Appendix E. Average Marginal Effect of Exogenous effects

Table E.5: Average Marginal Effects times standard deviation[†] of the exogenous variables

Exogenous Variables		PML- Λ					
		Professional (1)	Commercial (2)	Artisan (3)	Industrial		
					Builder (4)	Food/Agri (5)	Services (6)
Exogenous							
Age	direct	-0.14%	-0.54%	-0.69%	-0.50%	-0.45%	-0.38%
	indirect y	0.000%	-0.045%	0.007%	0.008%	0.008%	0.001%
	indirect $y^0 \neq y$	0.002%	-0.007%	0.008%	0.006%	0.009%	0.003%
n child	direct	0.38%	1.26%	1.28%	0.92%	1.06%	0.73%
	indirect y	0.003%	0.064%	-0.012%	-0.012%	-0.015%	-0.001%
	indirect $y^0 \neq y$	-0.001%	0.010%	-0.012%	-0.007%	-0.013%	-0.004%
n servant	direct	-0.87%	-2.38%	-2.54%	-1.70%	-2.18%	-1.42%
	indirect y	-0.004%	0.024%	-0.005%	-0.006%	-0.005%	0.000%
	indirect $y^0 \neq y$	-0.002%	0.005%	-0.003%	-0.003%	-0.002%	-0.001%
married [‡]	direct	-0.050%	-0.323%	-0.358%	-0.269%	-0.258%	-0.201%
	indirect y	0.003%	-0.041%	0.008%	0.008%	0.009%	0.001%
	indirect $y^0 \neq y$	0.002%	-0.008%	0.006%	0.004%	0.006%	0.002%
stayerp [‡]	direct	0.241%	0.445%	0.381%	0.273%	0.370%	0.229%
	indirect y	0.018%	-0.167%	0.025%	0.028%	0.030%	0.002%
	indirect $y^0 \neq y$	0.014%	-0.030%	0.025%	0.020%	0.025%	0.011%

Notes: [†] Average Marginal Effect stands for Average Marginal Effect times the standard deviation of exogenous variable, $AME_{k,y} = \hat{E}_i \left[\frac{d\hat{p}_{y,i}}{dx_{k,i}} \hat{s}d_i(x_{k,i}) \right]$. [‡] Assuming they are continuous. “direct” is the direct effect of the individual characteristic on the probability of choosing a certain occupation y, “indirect y” is the indirect effect of this characteristic affecting the beliefs on their neighbors following a certain occupation y, and “indirect $y^0 \neq y$ ” is the indirect effect of the characteristic affecting the beliefs on their neighbors following any other occupation $y^0 \neq y$ (see Appendix D).

References

- Aguirregabiria, V., Mira, P., 2007. Sequential estimation of dynamic discrete games. *Econometrica* 75 (1), 1–53.
- Akerlof, G. A., Kranton, R. E., 2000. Economics and identity. *Quarterly journal of Economics*, 715–753.
- Andersen, E. B., 1970. Asymptotic properties of conditional maximum-likelihood estimators. *Journal of the Royal Statistical Society. Series B (Methodological)*, 283–301.
- Anderson, S. P., De Palma, A., Thisse, J. F., 1992. *Discrete choice theory of product differentiation*. MIT press.
- Aradillas-Lopez, A., 2010. Semiparametric estimation of a simultaneous game with incomplete information. *Journal of Econometrics* 157 (2), 409–431.
- 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.
- Bertrand, M., Luttmer, E. F., Mullainathan, S., 2000. Network effects and welfare cultures. *The Quarterly Journal of Economics* 115 (3), 1019–1055.
- Bisin, A., Moro, A., Topa, G., 2011. The empirical content of models with multiple equilibria in economies with social interactions. Tech. rep., National Bureau of Economic Research.
- Blau, D. M., Robins, P. K., 1990. Job search outcomes for the employed and unemployed. *Journal of Political Economy*, 637–655.
- Blume, L. E., Brock, W. A., Durlauf, S. N., Ioannides, Y. M., 2010. *Handbook of Social Economics*. Vol. 1B. Cambridge: MIT Press, Ch. 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., 1994. Immigration, ethnic identity, and assimilation: The intergenerational transmission of immigrant skills. In: *Economic aspects of international migration*. Springer, pp. 139–154.
- 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: Blume, L., Durlauf, S. (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.
- Cappellari, L., Tatsiramos, K., 2015. With a little help from my friends? quality of social networks, job finding and job match quality. *European Economic Review* 78, 55–75.
- Chamberlain, G., 1980. Analysis of covariance with qualitative data. *Review of Economic Studies* 47 (1), 225–238.
- De Giorgi, G., Pellizzari, M., Redaelli, S., 2007. Be as careful of the books you read as of the company you keep: Evidence on peer effects in educational choices.
- De Paula, A., 2013. Econometric analysis of games with multiple equilibria. *Annu. Rev. Econ.* 5 (1), 107–131.
- Dowd, B. E., Greene, W. H., Norton, E. C., 2014. Computation of standard errors. *Health services research* 49 (2), 731–750.
- Duflo, E., Saez, E., 2003. The role of information and social interactions in retirement plan decisions: Evidence from a randomized experiment. *The Quarterly Journal of Economics* 118 (3), 815–842.
- Dyos, H. J., 1953. *Workmen's Fares in South London, 1860-1914*. University College of Leicester.
- Evans, W. N., Oates, W. E., Schwab, R. M., 1992. Measuring peer group effects: A study of teenage behavior. *Journal of Political Economy* 100 (5), 966–991.
- 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*, 166–170.
- Gaviria, A., Raphael, S., 2001. School-based peer effects and juvenile behavior. *Review of Economics and Statistics* 83 (2), 257–268.
- Glaeser, E. L., Sacerdote, B., Scheinkman, J. A., 1996. Crime and social interactions. *The Quarterly Journal of Economics* 111 (2), 507–548.
- Goldsmith-Pinkham, P., Imbens, G. W., 2013. Social networks and the identification of peer effects. *Journal of Business & Economic Statistics* 31 (3), 253–264.
- Graham, B., 2008. Identifying social interactions through conditional variance restrictions. *Econometrica* 76 (3), 643–660.
- Graham, B., Hahn, J., 2005. Identification and estimation of the linear-in-means model of social interactions. *Economics Letters* 88 (1), 1–6.
- 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, 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.
- Helmets, 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. Tech. rep., 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.
- Kramarz, F., Skans, O. N., 2007. With a little help from my... parents? family networks and youth labor market entry. Unpublished paper, CREST.
- Krauth, B. V., 2006. Simulation-based estimation of peer effects. *Journal of Econometrics* 133 (1), 243–271.
- Krinsky, I., Robb, A. L., 1986. On approximating the statistical properties of elasticities. *The Review of Economics and Statistics*, 715–719.
- Laschever, R., 2009. The doughboys network: Social interactions and the employment of world war i veterans.
- 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* 96 (3), 402–417.
- Lin, X., 2010. Identifying peer effects in student academic achievement by spatial autoregressive models with group unobservables. *Journal of Labor Economics* 28 (4), 825–860.
- Manski, C. F., 1993. Identification of endogenous social effects: The reflection problem. *The Review of Economic Studies* 60 (3), 531–542.
- McFadden, D. L., 1974. Conditional Logit Analysis of Qualitative Choice Behavior. in Paul Zarembka, ed. *Frontiers in Econometrics* (New York: Academic Press).
- Moffitt, R., 2001. Social dynamics. Cambridge: MIT Press, Ch. 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. Tech. rep., National Bureau of Economic Research.
- NAPP, 2015. Minnesota population center. north atlantic population project: Complete count microdata. version 2.2 [machine-readable database]. Minneapolis: Minnesota Population Center.
- Neyman, J., Scott, E. L., 1948. Consistent estimates based on partially consistent observations. *Econometrica: Journal of the Econometric Society*, 1–32.
- Patacchini, E., Zenou, Y., 2012. Ethnic networks and employment outcomes. *Regional Science and Urban Economics* 42 (6), 938–949.
- 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.
- Sacerdote, B., 2001. Peer effects with random assignment: Results for dartmouth roommates. *Quarterly Journal of Economics* 116 (2), 681–704.
- Schmiechen, J. A., 1984. Sweated industries and sweated labor: The London clothing trades, 1860-1914. Croom Helm London.
- Schürer, K., Woollard, M., 2003. National sample from the 1881 census of great britain [computer file]. Colchester, Essex: History Data Service, UK Data Archive [distributor].
- Smith, R. M., 1904. The religious life of London. Hodder and Stoughton.
- Soetevent, A. R., Kooreman, P., 2007. A discrete-choice model with social interactions: with an application to high school teen behavior. *Journal of Applied Econometrics* 22 (3), 599–624.
- 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.
- 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: Jess Benhabib, A. B., Jackson, M. O. (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.
- Trogdon, J., Nonnemaker, J., Pais, J., 2008. Peer effects in adolescent overweight. *Journal of Health Economics* 27 (5), 1388–1399.

- 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.
- Woollard, M., 1998. The classification of occupations in the 1881 census of england and wales. *History and Computing* 10 (1-3), 17–36.
- Zimmerman, D. J., 2003. Peer effects in academic outcomes: Evidence from a natural experiment. *The Review of Economics and statistics* 85 (1), 9–23.