

ORIGINAL ARTICLE

Open Access

Ethnic capital and self-employment: a spatially autoregressive network approach

Xingang Wang^{1*} and Sholeh A Maani²

* Correspondence:

xwan254@aucklanduni.ac.nz

¹Department of Economics, The University of Auckland, 12 Grafton Road, Auckland 1010, New Zealand
Full list of author information is available at the end of the article

Abstract

Immigrants experience substantial disadvantages in employment in the host country. “Ethnic capital” (e.g. the ethnic network) is argued to provide a niche for immigrants. Previous international studies adopt either ethnic concentration or language as proxy for immigrants’ network in host country. In this paper we introduce a new “spatial autoregressive network approach” to construct a dynamic network variable from micro-data to capture the effects of social and resource networks for immigrants. This approach allows each individual’s self-employment decision to be geographically and ethnically correlated with that of other individuals. We further show that there are three advantages of this approach: it captures a more accurate effect of networks; it provides a better estimation of the impact of other socio-economic variables on self-employment; and it provides a better data fit. The empirical findings of this study strongly suggest that ethnic capital plays a vital role in relation to immigrants’ self-employment decisions.

JEL classification codes: J23, J61, Z13

Keywords: Ethnic capital; Ethnic network; Ethnic concentration; Spatial approach; Immigrant; Self-employment

1. Introduction

Two of the labour market’s assimilation outcomes are of special interest to labour economists and policy makers: employment assimilation and wage assimilation of immigrants in their host country. Many studies have attempted to analyse the earnings assimilation process of immigrants, and they indicate that many variables affect assimilation outcomes. Self-employment is an important form of employment for immigrants, but it has not been extensively examined by previous economic studies. This paper examines the effect of “ethnic capital” (e.g. ethnic network and ethnic concentration) on immigrant choices to engage in self-employment as opposed to employment as an employee.

Self-employment provides immigrants with another niche for assimilation. There are several socio-economic factors that significantly influence immigrants’ self-employment decisions. Chiswick (1978) argued that, compared with natives, immigrants are disadvantaged in the host country’s labour market because they lack English-language skills, social networks, knowledge of local customs, information about job opportunities, transferability of skill and firm-specific training. For these reasons new immigrants whose first language is different from the language of the host country face barriers to finding a job. As such, it

may take a long time for their income to converge to the income level of comparable natives in the host country.

However, rather than seeking employment in the waged sector, immigrants can choose to operate their own business in order to avoid at least some of the disadvantages listed in the preceding paragraph. Thus, immigrants may be more inclined to choose self-employment. A growing number of international studies have found that there are an increasing number of immigrant-owned businesses in countries that traditionally accept immigrants. Light and Sanchez (1987) found evidence that immigrants have been more likely than natives to be self-employed for the last one hundred years. In this paper we focus on this important but less studied form of immigrant employment.

The literature on ethnic capital or networks generally adopts either ethnic concentration (e.g. Edin et al. 2003; Andersson and Hammarstedt 2011) or linguistic concentration (e.g. Bertrand et al. 2000) as the proxy for immigrants' network in the host country's labour market. In addition, prior empirical studies of immigrant earnings or employment have assumed that the labour market performance of an individual is independent and identically distributed (*i.i.d.*).

In this paper, we propose an approach based on the influence of self-employment decisions, typically made by other members of the ethnic network in a locality, on an immigrants' self-employment decisions during a specific year and location. This spatial approach, we believe, shows more clearly the effect of the quality of the ethnic network. We also relax the assumption of *i.i.d.* in the previous models by introducing a dynamic spatial autoregressive model of immigrant choices, and we examine the effect of this new approach on results. In addition, we incorporate a spatial approach to estimate the magnitude and significance of the network effect on immigrants' employment decisions. This is fundamentally different from conventional models. The spatial approach allows researchers to capture the correlation of self-employment choices by an ethnic group in a locality in each time period. Conventional models have at best controlled for national level or local average group characteristics. We argue that our modelling approach of a 'dynamic ethnic spatial' network variable demonstrates the network impact more clearly. This variable has been dubbed the 'weighted ethnic spatial lag' and is discussed in our model outline. This approach allows an individual's self-employment decision to be geographically and ethnically correlated with that of other individuals.

We hypothesize that "ethnic capital" is a key resource which provides assistance to immigrants in establishing their own businesses, which may be particularly helpful for immigrants whose first language differs from the language of the host country. We construct a dynamic "ethnic spatial network" variable (also known as "weighted ethnic spatial lag", (Wy), as specified in the model section of our paper (Section 3.1.2), from individual-level data to account for networks in order to capture the effects of social and resource networks for immigrant groups. We apply this analysis to a new and rich longitudinal data set on immigrants. This approach offers three advantages: (1) It provides a better estimation of the quality or the degree of association of immigrants' networks; (2) Compared to the conventional approach, the spatial approach provides a better data fit; (3) The spatial model also provides a better estimation of the impact of personal characteristics and human capital variables. We show in our paper (Section 3.1.2) that when we exclude this variable, other coefficients partly capture the network

effect, resulting in a bias. We show that in our analysis other socio-economic variables were particularly affected by omitting this variable.

This paper is organised as follows: Section 2 provides hypotheses based on ethnic capital. In Section 3 we discuss the econometric models. Section 4 describes the data set used in this study. Empirical analyses are provided in Section 5, followed by the conclusion in Section 6.

2. Hypotheses

2.1 Ethnic capital

The concept of ethnic capital in the context of immigration economics was first advanced by Borjas (1992). He observed that the second generation's labour market performance depends on the skill level of their father's generation and on the overall ethnic environment.

In this paper, we adopt a wider definition, where "ethnic capital" can also refer to the ethnic concentration and network of an immigrant group. We hypothesize that immigrants can find certain helpful features awaiting them in the host country, such as an existing/established network of earlier immigrants with shared ethnicity, which they can join; or a substantial number of earlier immigrants from their ethnic group who have settled in the location in which the new immigrant chooses to live. Such features are known collectively as "ethnic capital". In other words, ethnic capital is the inherent trust and advantages which stem from, and belong to, a certain ethnic/cultural group.

2.1.1 Ethnic network

Ethnic network links immigrants, and it works as a platform to distribute economic resources. Immigrants may face many disadvantages, such as the lack of financial resources; lack of information about the local market and regulations; and less proficiency in the local language.

Van Auken and Neeley (1998), Anthony (1999) and Lofstrom (2002) noted that the ethnic network helps immigrants to obtain sufficient start-up financial capital. This assistance may include matching immigrant businesses' demand for liquidity, with financial resources from the ethnic enclave, and by providing necessary labour force and management. In addition, the ethnic network can work as an informal, instead of formal, financial sector to provide funds for immigrants (Bond and Townsend 1996). By studying the case in the United States, they indicated that this kind of informal sector is much more efficient than the formal financial sector, and immigrant entrepreneurs prefer to seek financial resources from the informal sector in preference to the formal. For example, the informal sector lowers the costs of information, searching, and monitoring for immigrants. These effects generate greater opportunities that help immigrants perform better economically through broader economic mechanisms. As a result, it is expected that ethnic networks increase the propensity of immigrant entrepreneurship in the host country (e.g. Borjas 1995; Toussaint-Comeau 2008).

The ethnic network not only provides funds to potential immigrant entrepreneurs, but it also provides support through culture and tradition. Under well-established cultures, some ethnic groups, such as the Gujarati Indians, have a tradition of entrepreneurship

(Aldrich et al. 1984). They share the common beliefs of entrepreneurs and assist potential entrepreneurs to set up businesses through their ethnic network. In addition, the ethnic network also promotes business communication and development for immigrant-owned businesses within the ethnic enclave. In this regard, when Wilson and Martin (1980) examined the case for the United States, they found that Cuban-owned firms were most likely to have Cuban suppliers; Rajman and Tienda (2003) found that Korean-owned and Mexican companies also preferred to do business with companies owned by other immigrants in the United States.

2.1.2 Ethnic concentration

It is increasingly recognized that immigrants are potentially both complements and substitutes for each other. In the broader sense of the term, complementary mechanisms (e.g. job creation, provision of financial resources, and demand for ethnic products) positively assist immigrants to be self-employed; while a substitution mechanism (e.g. as employees competing for jobs, or competition in business) negatively influences immigrants' labour market performance. As a result, immigrants benefit from concentrating geographically by ethnicity once the complementary mechanisms dominate the substitution mechanisms. We discuss the empirical evidence on these two opposing effects below. The question of the dominant effect remains to be an empirical question across different immigrant groups or settings.

Complementary mechanisms Immigration influences both the host country's labour market and the goods market (Jean and Jimenez 2011). Immigration provides both labour supply and demand for goods and services in the host country. Because of their ethnic background, immigrants may prefer products and services which could not be satisfied by local suppliers. However, immigrants themselves can fill this gap sufficiently and service other immigrants, probably from the same ethnic group (Waldinger 1986). This factor is closely related to the geographical location variable, as location implies both opportunities to set up businesses, and it also reflects the size of the potential target market.

"Ethnic enclave" is often defined as the "concentration of immigrants in a residential location" (Borjas 1986). Ethnic enclaves provide many resources for immigrants including a larger and potentially cheaper labour force, ethnic solidarity, vertical integration and a protected market (Aguilera 2009). An ethnic enclave provides a market, and it reduces the barriers to employment by establishing businesses for and by immigrants. For example, Rajman and Tienda (2003) point out that by being employed in an ethnic enclave, immigrants are offered more opportunities for training in entrepreneurship that qualify them for self-employment. Most previous international studies indicated that the ethnic enclave (ethnic concentration) has positive (complementary) influences on immigrants' self-employment decisions (e.g. Wilson and Portes 1980; Borjas 1986; Toussaint-Comeau 2008; Le 2000).

Again, the immigrant market is noted to be a "protected market" because immigrants have a specific demand for ethnic goods and services (e.g. Boyd 1990; Aldrich et al. 1985; Aldrich and Waldinger 1990; Aguilera 2009). This specific demand for ethnic products and services from the immigrant community increases with the size of the immigrant population of that ethnic group in a specific region. In addition,

immigrant entrepreneurs are more efficient in serving this kind of ethnic-oriented demand as they know their ethnic immigrants' preferences, demand, culture, norms, and customs better than the local businesses. Therefore, with a larger ethnic enclave, more business opportunities can be generated for potential immigrant entrepreneurs from the "protected market". Furthermore, immigrants share social and economic capital through their network. Therefore, in this case, ethnic concentration works positively and promotes self-employment among immigrants.

We hypothesize that ethnicity is a kind of capital for immigrants. In our hypothesis section, we explore that when immigrants concentrate locally, it generates greater socio-economic resources (including financial assistance), information and demand from other immigrants, which facilitates immigrants' self-employment.

Substitution mechanisms Some international empirical studies have observed a negative relationship between ethnic concentration and the propensity for new self-employment among immigrants (e.g. Clark and Drinkwater 1998, 2000; Aldrich and Waldinger 1990). Aldrich and Waldinger (1990) claimed that the negative effect of the ethnic enclave on immigrants' entrepreneurship is due to the "effect of limiting entrepreneurial opportunities" and to the existence of too much competition within the ethnic enclave. In this case, the growing ethnic enclave could not generate sufficient opportunities and other socio-economic resources for promoting immigrant self-employment.

In addition, with the growth of the ethnic enclave, the immigrant market becomes a non-neglected market in the host country, leading local businesses to hire immigrants to serve and develop the immigrant market. In this scenario, there will be increased job opportunities as employees offered to immigrants in the mainstream economy, which also decreases the propensity for self-employment among immigrants. As a result, if the substitution mechanisms with greater waged opportunities dominate the complementary mechanisms, ethnic concentration decreases the propensity for immigrant entrepreneurship.

3. Model

3.1 Measurement of ethnic network

In order to explore the effects of ethnic networks, previous empirical studies have adopted either ethnic concentration (e.g. Aguilera 2009; Damm 2009; Edin et al. 2003; Toussaint-Comeau 2008; Borjas 1995; Andersson and Hammarstedt 2011) or linguistic concentration (e.g. Bertrand et al. 2000) as the proxy for an ethnic network. Unlike these studies, we adopt the "spatial approach" to account for ethnic networks and concentration in order to capture the effects of social and resource networks for immigrant groups. We added a weighted ethnic spatial lag variable and compare this to the conventional model.

3.1.1 Weighted ethnic spatial lag

We constructed an ethnic spatial network variable – 'weighted ethnic spatial lag' as the proxy of immigrants' ethnic network to represent the individual's network of economic resources, in addition to ethnic concentration. By doing so, we are able to separate the network-specific resource effect from the more general ethnic concentration.

We hypothesize that both ethnic networks and ethnic concentration influence immigrants' self-employment decisions.

W is a $n \times n$ ethnic spatial weighted matrix, which shows the first-order ethnic and geographical (ethnic-spatial) relationship among individuals. Before the discussion of W , the first-order ethnic spatial neighbourhood matrix E will be introduced by an example. Suppose individuals P1, P2, P4 and P6 are all from Asia; P1 and P4 are all located in region A, while individuals P2 and P6 are located in region B. P3, P5 and P7 are from Europe, all of them are located in region B. Thus, the 7×7 first-order ethnic-spatial neighbourhood matrix E is:

$$E = \begin{pmatrix} & P1 & P2 & P3 & P4 & P5 & P6 & P7 \\ P1 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ P2 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ P3 & 0 & 0 & 0 & 0 & 1 & 0 & 1 \\ P4 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ P5 & 0 & 0 & 1 & 0 & 0 & 0 & 1 \\ P6 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ P7 & 0 & 0 & 1 & 0 & 1 & 0 & 0 \end{pmatrix} \quad (1)$$

When the elements of matrix E are zeros, individuals are not deemed to be first order ethnic-spatial neighbours. In addition, the diagonal elements of the above matrix are zeros which means individuals are not considered as neighbours to themselves.

In order to define a "weighted ethnic spatial lag", the ethnic spatial matrix E should be normalised by unifying the row sums, such that we can form the ethnic spatial weighted matrix W :

$$W = \begin{pmatrix} & P1 & P2 & P3 & P4 & P5 & P6 & P7 \\ P1 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ P2 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ P3 & 0 & 0 & 0 & 0 & 1/2 & 0 & 1/2 \\ P4 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ P5 & 0 & 0 & 1/2 & 0 & 0 & 0 & 1/2 \\ P6 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ P7 & 0 & 0 & 1/2 & 0 & 1/2 & 0 & 0 \end{pmatrix} \quad (2)$$

In our study, the "ethnic-spatial weighted matrix" is considered to be dynamic. Our matrix is constructed with micro data, based on the three conditions of¹: 1) ethnic group, 2) region of residence, and 3) year of survey. Therefore, if immigrants were to shift location, the demographic composition of regions would alter (for example, the number of Chinese immigrants in a specific region would change). Our ethnic spatially weighted matrix, W , captures this dynamic aspect. In addition, since we derive W by normalizing the ethnic-spatial neighbourhood matrix, 'E', through unifying the row sums, the elements of W always fall in the range of 0 and 1. As such, the row sum reflects for a typical immigrant, how many immigrants (from the same ethnic group) are living in the same place with that immigrant. Therefore, when the number of immigrants from a typical ethnic group in a specific region changes, the relative elements of W will be changed as well.

We are interested in better-understanding the migrants' network effect through the data. The spatial model provides a new and more relevant theoretical and empirical framework to investigate the effect of ethnic capital.

3.1.2 Ethnic spatial model

The data generating process for the situation when the value of one observation i depends on the value of its neighbour j 's observation (e.g. LeSage and Pace 2009) is as below:

$$y_i = \alpha_i y_j + \beta X_i + \varepsilon_i \quad (3)$$

$$y_j = \alpha_j y_i + \beta X_j + \varepsilon_j \quad (4)$$

$$\varepsilon_i \sim N(0, \sigma^2), \quad \varepsilon_j \sim N(0, \sigma^2)$$

Thus, equations (3) and (4) imply a “simultaneous data generating process” that shows the dependence of y_i and y_j and vice versa. This analytical feature leads us to a data generating process which is an “Ethnic Spatial Autoregressive Process”, and the following expression:

$$y_i = \rho \sum_{j=1}^n W_{ij} y_j + \beta X_i + \varepsilon_i \quad (5)$$

$$\varepsilon_i \sim N(0, \sigma^2) \quad i = 1, \dots, n$$

where X_i is a vector of socio-economic variables for individual i . In our analysis, y_i and y_j represent self-employment choices by individuals i and j . The “ethnic neighbour” is defined as individuals who are from the same ethnic group and in the same location. Thus, $\sum_{j=1}^n W_{ij} y_j$ is the “weighted ethnic spatial lag” in this context, and it represents the linear combination of individual i 's ethnic neighbours' self-employment choices.

As a result, the matrix version of equation (5) is:

$$y = \rho W y + \beta X + \varepsilon \quad (6)$$

$$\varepsilon \sim N(0, \sigma^2 I_n)$$

where $N(0, \sigma^2 I_n)$ represents the zero mean disturbances process with the constant variance σ^2 . I_n is the n -dimensional identity matrix.

Under the ethnic capital hypothesis, individuals' incomes depend on ethnic capital and other socio-economic variables. In this setting, one can define individuals who are from the same ethnic group and location as first-order “ethnic neighbours”. Thus, “weighted ethnic spatial lag” represents the case where an individual's labour market performance is influenced by its ethnic neighbours' labour market performance and other ethnic-capital factors in that location. Therefore, the matrix version of our model is:

$$y = \alpha l_n + \rho W y + \tau EC + \beta X + \varepsilon \quad (7)$$

where y is immigrants' economic performance (e.g. self-employment outcome), X is a vector of socio-economic variables; $W y$ is the weighted ethnic spatial lag vector which indicates the first-order ethnic-spatial relationship among individuals; l_n is a vector of ‘ones’ and associated with the parameters α and β . Thus, the coefficient ρ indicates the size of the effect of the network in a specific region. Furthermore, we are also able to construct a full social networking variable for that individual via W ; and it captures all the information of a network.

Ethnic concentration (EC) has been defined in various forms across studies. For example, Borjas (1986) argued that a Hispanic enclave definitely helped the Hispanic

immigrant entrepreneurs in the United States due to the cultural and language similarities for three Hispanic groups (Mexicans, Cubans, and other Hispanics). He defined the ethnic concentration variable as the proportion of Hispanic population of the MSA's² population in the United States.

In this paper, we adopted a similar approach to Borjas' (1986) with

$$EC_{kl} = \frac{Population_{kl}}{Population_l} \quad (8)$$

where "k" denotes ethnic group, and "l" represents a specific state or region³.

Due to the hypothesis of ethnic enclave, we would expect that the coefficient of the immigrant population size in a specific region should be positive in most cases⁴.

From rearranging equation (7) we derive:

$$\begin{aligned} (I_n - \rho W)y &= \alpha l_n + \tau EC + \beta X + \varepsilon \quad (9) \\ y &= (I_n - \rho W)^{-1} \alpha l_n + (I_n - \rho W)^{-1} \tau EC + (I_n - \rho W)^{-1} \beta X + (I_n - \rho W)^{-1} \varepsilon \\ \varepsilon &\sim N(0, \sigma^2 I_n) \end{aligned}$$

In comparison, the previous basic econometric model for immigrants' labour market performance in matrix version is:

$$y = \alpha l_n + \gamma X + \varepsilon \quad (10)$$

It is noteworthy that the coefficient β for the explanatory variables in equations (7), and (9), where we have included the "weighted ethnic spatial lag" in our model, is different from the coefficient γ in the conventional model (equation 10). Therefore when we take the network effect into account, all estimated coefficients need to be accordingly adjusted by the spatial dependence. As a result, the spatial model provides a better estimation of the effects of immigrants' personal characteristics and other socio-economic factors when the network effect is present, compared to the conventional model.

3.2 Spatially autoregressive discrete choice model

Following from our hypotheses of ethnic capital, we investigate how ethnic network influences immigrants' self-employment decisions. The logit model is widely employed in testing such discrete choices, as it approaches the random utility assumption to the self-employment choices (to be self-employed or not). In this study, we have adopted similar settings for a binary weighted spatial lag model as those of Adjemian et al. (2010).

Immigrant i chooses a form of employment (either to be self-employed (*S.E.*) or employed in the wage/salary (*W.S.*) sector) which will maximise his/her utility. For self-employment choice (*S.E.*) the utility for a recent male immigrant is given by:

$$U_i^{S.E.} = V_i^{S.E.} + \varepsilon_i^{S.E.} \quad (11)$$

where $V_i^{S.E.}$ shows the deterministic portion of utility, $\varepsilon_i^{S.E.}$ represents a random component. Then, the deterministic utility is composed of a set of explanatory variables and a weighted ethnic spatial lag (which represent social network effect):

$$V_i^{S.E.} = \beta' x_i + \rho' Wf(V_i^{S.E.}) \quad (12)$$

where x_i is a set of social-economic variables of immigrant i , such as educational attainment, years since migration, age, and other demographic variables and local

characteristics; and W is the spatial weight matrix which indicates the first order ethnic neighbourhood for every immigrant. As a result, the coefficient of W indicates the correlation of utility from choosing self-employment for all immigrants who are members of a particular ethnic network.

The immigrant makes a decision regarding which sector to be employed in: the self-employment ($S.E.$) sector or wage/salary ($W.S.$) sector. As a result, the decision rule for immigrant i is expressed as:

$$\begin{aligned}
 \Pr[y = S.E.] &= \Pr[U_i^{S.E.} > U_i^{W.S.}] \\
 &= \Pr[U_i^{W.S.} - U_i^{S.E.} < 0] \\
 &= \Pr[V_i^{W.S.} + \varepsilon_i^{W.S.} - V_i^{S.E.} - \varepsilon_i^{S.E.} < 0] \\
 &= \int I(\varepsilon_i^{W.S.} - \varepsilon_i^{S.E.} < V_i^{S.E.} - V_i^{W.S.}) f(\varepsilon_i^{W.S.}) d\varepsilon_i^{W.S.}
 \end{aligned}
 \tag{13}$$

where the indicator function I takes the value of one if the expression in parentheses is true, otherwise it is equal to zero. In addition, the independent random error assumption is held, and ε is identically Bernoulli distributed for all immigrants (see Adjemian et al. 2010). Therefore, the probability of immigrant i deciding to choose self-employment $S.E.$ is given by the logistic probability:

$$P_i^{S.E.} = \frac{1}{1 + \exp(V_i^{S.E.})}
 \tag{14}$$

In previous studies of spatial discrete choice models the network effect is treated as a signal or a kind of knowledge (see Goetzke, 2008; and Adjemian, et al, 2010), which means the spatial spill-over could be unidirectional but not multidirectional (e.g. Adjemian et al. 2010). Goetzke (2008) made the assumption following Anselin (2002) on transport choice modes that, “The model is conditional upon the observed neighbouring mode choices, which means that the spatial spill-over process is not modelled as an endogenous process. The advantage is that the estimation of this model type is straightforward to estimate”. For example, in this setting, once individual a has made a choice, individual b will learn this information and possesses this knowledge as one of the factors assisting him/her to make a decision. However, individual b 's decision cannot go back to influence individual a 's decision in the same round. Adjemian et al. (2010) also treated the weighted spatial lag variable as exogenous due to the nature of car purchases, and transactions costs which constrain a household's car purchases to be fixed in the short term. They note that, “As a result, spatial spill-overs in auto choices are necessarily unidirectional”. One could apply a similar logic to immigrant self-employment since moving across localities and engagement in self-employment takes time, and there are significant transaction costs.

However, the potential endogeneity of the weighted ethnic spatial lag variable in general (e.g. Goetzke and Andrade 2009), and for immigrant self-employment due to either multi-lateral (as opposed to unilateral) network effects, or the existence of potential unobservable variables that correlate with the spatial lag variable cannot be ruled out. Therefore, in this paper, we relax this assumption, such that the spatial spill-over could also be multidirectional rather than unidirectional⁵. We adopt the modelling approach that controls for potential endogeneity of the weighted ethnic spatial lag term

(Anselin 1990; Kelejian and Prucha 1998; and Kelejian and Robinson 1993), based on a Spatial Two Staged Least Squares. Hence, the weighted ethnic spatial lag term (ethnic network in this study) is treated as an endogenous variable⁶.

Anselin (1990), Kelejian and Prucha (1998), and Kelejian and Robinson (1993) proposed a Spatial Two-Stage Least Squares (2SLS) approach to control for the endogeneity of spatial lag. As Kelejian and Robinson (1993) have illustrated, all exogenous variables X , their spatial lag WX and higher spatial lags (e.g. W^2X , W^3X , ..., W^nX) work jointly as a set of instruments for the endogenous spatial lag Wy . In this study, as noted by Anselin (1999), according to the computational complexity of using the full set of instruments for the endogenous spatial lag, we have selected the first order spatial lag of all exogenous variables WX as well as X as the set of instruments. We find that this specification results in significantly better performance than the model that assumes the exogeneity of the weighted ethnic spatial lag variable, Wy .

3.2.1 Quality and strength of ethnic network

The effect of the quality (strength) of networks on immigrants' self-employment decisions is less examined.

Earlier advances in conventional models such as Bertrand et al. (2000), and Edin et al. (2003) measured ethnic capital as the interaction between ethnic group, neighbourhood and year. Among the studies on self-employment of immigrants, Sousa (2013) measured the quality of the local community based on human capital; and Toussaint-Comeau (2008) measured the quality of the network by "the average relative self-employment rate of the group in the U.S."

As noted earlier, in this paper we propose a spatial approach that captures the quality and strength of ethnic group self-employment choices. The 'weighted ethnic spatial lag' (ethnic network variable) does not simply reflect the interaction of ethnic group, neighbourhood and year, or average resources. Instead, in this approach the model includes the influence from self-employment decisions typically made by other members of the ethnic network in a locality and during a specific year. We believe that this approach more closely shows the correlation of individuals' decisions and also its effect on the decisions of the other group members. In addition, in the approach adopted we view the data as *i.i.d.* This is fundamentally different from usual linear models. As we show in our empirical results (Section 5), spatial models are preferred statistically in every case.

3.3 Review of the literature (other variables of interest)

Economists and sociologists have observed about six key determinants of immigrant entrepreneurship (refer to Le 2000; Evans 1989; Kidd 1993): educational attainment, labour market experience, economic requirements, marital status, industry and occupation factors, and the host country's language and ethnicity factors.

In the literature on self-employment, educational attainment is noted to have a significant influence resulting from two opposing forces (e.g. Le 2000). On the one hand, educational attainment reflects the ability of the individual, in particular, his or her managerial ability, to operate a business. On the other hand, individuals with higher educational attainment are less likely to be self-employed since education enhances the

propensity for a person to find employment in the waged sector. Therefore, the dominant impact is an empirical question.

Experience is argued to be either a “stock” (Evans 1989) or “flow” (Kidd 1993). In addition, labour market experience can be viewed as the accumulation of skills and market information. With greater experience, an individual will be more confident about operating a business. Secondly, age increases at the same time as an individual’s labour market experience increases. With the increase in age, personal learning capacity and the present value of future returns diminishes, so increasing age also decreases the propensity for self-employment.

Previous studies (e.g. Bernhardt 1994; Kidd 1993) have paid attention to the importance of economic requirements for entrepreneurship decision. For example, Kidd (1993) used age as a proxy for financial capital and adopted a binary variable, “rent”, to study immigrants’ propensity for self-employment. Kidd concluded that those who own their residence are more likely to select self-employment than those who rent a house.

Marital status is an indicator of stability, which thus provides implications for and background to a risky self-employed status. Borjas (1986) noted that married individuals are more likely to choose self-employment because married couples may like to “put up” or join financial resources to run their business. In addition, given family support, it would also reduce the unwillingness to take risks that an individual might face. As a result, marriage makes self-employment more feasible for an individual.

Since the first wave of the data was conducted six months after new immigrants settled in New Zealand, the variables such as proficiency in English, children, marriage, skill level, overseas self-employment experience and own dwelling in our model are treated as exogenous variables, as by design we incorporate only the first wave’s data for those variables.

It is also hypothesized that self-employment is partially affected by occupational status. According to the Middleman Minority Theory, the employment status of an individual is decided by the work undertaken (Bonacich and Modell 1980). Current employment provides work experience and training for potential entrepreneurs before they set up their own business. This is also a complementary explanation for the impact of experience, as more information about the market, business networks and business skills will be acquired during that period. Evans (1989) observed that individuals with a high occupational status are more likely to choose self-employment. More specifically, Le (2000) claimed that trade, sales, and managerial occupations require more relevant knowledge, and they also make self-employment more feasible.

The effect of skill in the host country’s language is significant and unambiguous. Host country’s language proficiency (e.g. English) reflects the immigrant’s integration into the general community. However, the effect of English-language skills in relation to self-employment is ambiguous, and it may vary by country, data, and cohort. On the one hand, a lack of skill in the host country’s language will hinder business communication with the native mainstream economy (e.g. Le 2000). On the other hand, a lack of English proficiency can increase the propensity for self-employment by satisfying the demand from other immigrants from the same ethnic group (e.g. Evans 1989). In addition, a third point of view is based on Disadvantage Theory (Light 1979): communication disadvantages make it difficult for immigrants to be employed in the wage sector; however, the same disadvantages encourage them to be self-employed.

Previous New Zealand studies (e.g. Poot 1998; Maré and Stillman 2009) have analysed the effect of human capital and personal characteristics on immigrants' labour market performance. However, the effect of ethnic capital (e.g. ethnic network and ethnic concentration) on immigrants' economic performance (especially self-employment outcome) in this context remains unknown.

In this paper we account for these factors in addition to our new network variable of interest.

4. Data

In this paper, we employ a new and rich panel dataset (Longitudinal Immigrant Survey: New Zealand (LisNZ)). In addition, we derived ethnic concentration variables from the published 2001 and 2006 population census tables and incorporated these to the LisNZ dataset to augment our data on ethnic concentration.

New Zealand is traditionally a country of immigrants. New Zealand has a considerable population of self-employed immigrants. Data from the 2006 census show the rate of self-employment among employed immigrants in New Zealand was about 15% (Nana and Sanderson 2008). This rate is higher than that for other traditionally immigrant-receiving countries. For example, the rate of self-employment among immigrants was 10.4% in Canada and 7.3% in the United States (Organisation for Economic Co-Operation and Development 2001).

The LisNZ project includes three interviews ("waves") with the same group of randomly selected immigrants⁷. Immigrants were sampled at the time they were granted residence. The survey sample was selected from migrants who were approved for permanent residence from 1 November 2004 to 31 October 2005. The first wave (Wave 1) interviews were conducted six months after new immigrants settled in New Zealand.

The second wave (Wave 2) is conducted 12 months after the Wave 1 survey. The last survey (Wave 3) was made 36 months after the immigrants had settled. The Wave 1 interview was conducted between 1st May 2005 and 30th April 2007. LisNZ survey includes a question regarding whether or not interviewees are self-employed. We have defined individuals as such where they have answered, "Yes".

In addition, like the New Zealand census, LisNZ classifies New Zealand into seventeen regions at regional council level. These regions are generally similar to the size of Metropolitan Statistical Areas (MSA) in the United States in terms of population and geographic spread⁸. LisNZ also provides detailed information on country of origin. As a result, it is possible to analyse specific ethnic group effects on ethnic capital. It also allows for the consideration of the specific effects for immigrants from different regions of Asia (notably Chinese and Indian immigrants), and allows comparison with other groups of immigrants (in particular, immigrants for the UK and Ireland)⁹.

4.1. Demographic characteristics

As discussed in the introduction, we are more interested in the effect of ethnic capital on immigrant choices to engage in self-employment as opposed to working for others. Therefore, we follow the setting adopted in recent studies to focus on the employed immigrant population (e.g. Bradley 2004; Le 2000; Lofstrom 2002). As a result, we have selected employed male immigrants aged between 20 and 55 years old¹⁰. We have dropped some observations due to missing data issues; the total sample size is 6,735.

The data set reflects a diverse group of immigrants in our sample. There are 1,587 observations contributed for immigrants from the United Kingdom and Ireland; 913 are Indian immigrants; and 781 are Chinese immigrants.

The average age of recently approved permanent residents from the United Kingdom and Ireland is the oldest among all five ethnic groups (38.45 years old) and is also higher than the sample average (36.11 years old). Recently approved permanent residents from China are generally the youngest among all ethnic groups, with an average age of 32.5 years.

Although the Chinese have the lowest average age, they tend to have a considerable number of New Zealand living experience years (“potential years since migration”). The mean “potential years since migration” for them is around 7.39 years, which is the highest among all five ethnic groups. In addition, the same figure for immigrants from the United Kingdom and Ireland is relatively low at about 6.66 years. Indian immigrants to New Zealand have the lowest New Zealand experience years among all ethnic groups (5.2 years). Therefore, in this sample, age and New Zealand living experiences are not significantly positively correlated.

Immigrants from the United Kingdom tend to have more experience in the labour market (18.0 years) than immigrants from other ethnic groups. Chinese (11.3 years) and Indian (11.7 years) immigrants to New Zealand have significantly fewer labour market experience years than immigrants from the United Kingdom and Ireland. As Additional file 1: Table S1 shows Chinese and Indian immigrants are more likely to have achieved higher educational attainment, and they are also younger than other immigrant groups. Therefore, it is reasonable that their average labour market experience is also the lowest among all five ethnic groups. Chinese immigrants have a relatively low rate of house/flat ownership among all ethnic groups, but their rate of self-employment is the highest. At the same time, while more than half of the immigrants from the United Kingdom and Ireland own homes, their rate of self-employment in this group is lower than for the Chinese immigrant group. The figure of approximately 13% for self-employment among immigrants from the United Kingdom and Ireland is slightly higher than the average of the sample.

5. Empirical results

5.1 General case

Table 1 provides all variable descriptions. Table 2 shows the coefficients¹¹ (odds ratios) obtained from the logit regressions for immigrants’ self-employment decision models, and Table 3 shows the average marginal effects for those variables. In this study, we have estimated the effect of ethnic network by two different approaches: i) conventional, and ii) spatial, and then we compare the results from these two approaches. The spatial model incorporates the potential endogeneity of the weighted ethnic spatial lag (network variable).

We use the “Panel Logit Model” to estimate the effect of human capital and ethnic capital on immigrants’ self-employment decisions¹². This estimation process is suitable for the data, and it allows us to exploit the panel feature of the data¹³.

In Table 2, the odds ratios for the two conventional models, i.e., columns (1) and (2) show results from conventional models without controlling for spatial dependence. The spatial models of interest (columns (3) and (4) in Table 2) explore the effect of an ethnic network on immigrants’ self-employment decisions with a spatial approach.

Table 1 Variable list and definitions

Human capital	
Labour market experience (Experience)	This is a derived variable which is equal to an individual's current age minus the age of graduation.
Proficiency in English*	Binary variable=1 proficient in English (self-reported).
Overseas self-employment experience	This variable represents the years of overseas self-employment experience of immigrants.
High skilled*	Binary variable=1 if that individual's highest education attainment is Bachelor or higher.
Personal characteristics	
Years Since Migration (YSM)	This variable represents the duration of immigration.
Married*	Binary variable=1 if married.
Children*	Binary variable=1 if has children.
Own dwelling*	Binary variable=1 if owns house/flat.
Manager & professional*	Binary variable=1 if occupation is manager or professional.
Ethnic capital	
Ethnic network	The weighted ethnic spatial lag (Wy. Refer to Section 3.1.2).
Ethnic concentration	The lag of proportion of the population of a specific ethnic group to the total population size in a specific region (EC_{ik} . Refer to Section 3.1.2)

Note: *These variables are based on the first wave data.

Results in the second column, in comparison, show the effect of an ethnic network through the conventional approach (the variable of ethnic concentration serves as the proxy for immigrants' network in the host country).

Based on our earlier discussion, we expect the ethnic concentration variable to reflect the outcome of a substitution and complementary mechanisms for immigrants in our spatial model. That is, once weighted ethnic spatial lag has been controlled for, ethnic concentration shows the dominant result of the complimentary and substitution mechanisms for immigrants. As such, the spatial model (column (3)) shows the effect of spatial ethnic network effects, and the second spatial model (column (4)) reflects the joint effects of network and ethnic concentration by controlling for the correlation of self-employment choices by immigrants' ethnic network.

As Tables 2 and 3 show, immigrants' network effect is positively correlated with recent male immigrants' self-employment decisions¹⁴. Ethnic network has a strong significant positive average marginal effect on the propensity of self-employment for immigrants. As a result, immigrants' self-employment decisions are positively correlated with each other, which means the immigrants' network promotes entrepreneurship among recent immigrants in New Zealand.

Compared to the conventional approach, the spatial approach offers three advantages:

Firstly, it captures a much more accurate effect of networks. The coefficient of weighted ethnic spatial lag (ethnic network) is highly significant, which indicates that immigrants' self-employment decisions remain correlated, ethnically and spatially, after controlling for other socio-economic variables. A positive coefficient of weighted ethnic spatial lag indicates ethnic network plays a positive role in relation to ethnic entrepreneurship. However, the conventional approach fails to capture that significant effect. In the second conventional model (column (2) in Tables 2 and 3), the

Table 2 Estimates of self-employment decision with network effect (odds ratios)

	Conventional		Spatial	
	(1)	(2)	(3)	(4)
Ethnic capital				
Network effect†	/	/	18.55***	18.48***
	/	/	(1.93)	(1.93)
Ethnic concentration effect	/	0.08	/	0.04
	/	(0.06)	/	(0.06)
Human capital				
Experience	0.30***	0.30***	0.36***	0.36***
	(0.03)	(0.03)	(0.03)	(0.03)
Experience -squared	-0.005***	-0.005***	-0.007***	-0.007***
	(0.001)	(0.001)	(0.001)	(0.001)
Self-employment experience (origin)	0.20***	0.20***	0.20***	0.20***
	(0.02)	(0.02)	(0.02)	(0.02)
Self-employment experience-squared (origin)	-0.01***	-0.01***	-0.01***	-0.01***
	(0.001)	(0.001)	(0.001)	(0.001)
High skilled	-0.12	-0.12	-0.66***	-0.66***
	(0.1)	(0.1)	(0.11)	(0.11)
English proficiency	-2.16***	-2.14***	-0.72***	-0.71***
	(0.13)	(0.13)	(0.20)	(0.20)
Manager & professional	1.61***	1.61***	1.33***	1.34***
	(0.1)	(0.1)	(0.1)	(0.1)
Personal characteristics				
Years Since Migration (YSM)	0.42***	0.42***	0.23***	0.23***
	(0.03)	(0.03)	(0.03)	(0.03)
YSM-squared	-0.01***	-0.01***	-0.01***	-0.01***
	(0.001)	(0.001)	(0.001)	(0.001)
Married	0.66***	0.66***	0.44***	0.44***
	(0.07)	(0.07)	(0.07)	(0.07)
Children	-0.42***	-0.43***	-0.46***	-0.46***
	(0.1)	(0.1)	(0.1)	(0.1)
Own dwelling	0.33***	0.33***	0.22***	0.22***
	(0.1)	(0.1)	(0.1)	(0.1)
Observations	6735	6735	6735	6735
Log likelihood	-11635.10	-11634.39	-11593.64	-11593.42
Wald Chi(2)	1268.03	1269.25	1370.93	1371.00
AIC	23294.20	23294.78	23215.28	23214.84

Note: †Network (weighted ethnic spatial lag 'Wy') is treated as an endogenous variable.

(1) Panel Logit estimations.

(2) Standard errors in parentheses.

(3) *p<0.10 **p<0.05 ***p<0.01.

coefficient of ethnic concentration (which was adopted as the proxy for ethnic network in prior studies) is not statistically significant. Therefore, by adopting the conventional approach, the significant positive effect of ethnic network cannot be observed; and we may mistakenly conclude that ethnic network does not matter for ethnic entrepreneurship.

Table 3 Estimates of self-employment decision with network effect (average marginal effects)

	Conventional		Spatial	
	(1)	(2)	(3)	(4)
Ethnic capital				
Network effect†	/	/	0.075***	0.075***
	/	/	(0.001)	(0.001)
Ethnic concentration effect	/	0.0003	/	0.0002
	/	(0.0002)	/	(0.0002)
Selected human capital				
Experience	0.0012***	0.0012***	0.0015***	0.0015***
	(0.0001)	(0.0001)	(0.0001)	(0.0001)
Self-employment experience (origin)	0.0008***	0.0008***	0.0009***	0.0009***
	(0.001)	(0.0001)	(0.0001)	(0.0001)
High skilled	-0.0005	-0.0005	-0.003***	-0.003***
	(0.0004)	(0.0004)	(0.0005)	(0.0005)
English proficiency	-0.0258***	-0.0252***	-0.004***	-0.004***
	(0.0037)	(0.0037)	(0.002)	(0.002)
Manager & professional	0.0066***	0.0066***	0.005***	0.005***
	(0.0006)	(0.0006)	(0.0005)	(0.0005)
Selected personal characteristics				
Years Since Migration (YSM)	0.0017***	0.0017***	0.0009***	0.0009***
	(0.0001)	(0.0001)	(0.0001)	(0.0001)
Married	0.0028***	0.0028***	0.0018***	0.0018***
	(0.0004)	(0.0004)	(0.0003)	(0.0003)
Children	-0.0017***	-0.0017***	-0.0018***	-0.0018***
	(0.0004)	(0.0004)	(0.0004)	(0.0004)
Own dwelling	0.0015***	0.0014***	0.0013***	0.0013***
	(0.0005)	(0.0005)	(0.0004)	(0.0004)

Note: †Network (weighted ethnic spatial lag 'W_y') is treated as an endogenous variable.

(1) Panel Logit estimations.

(2) Standard errors in parentheses.

(3) *p<0.10 **p<0.05 ***p<0.01.

Secondly, as discussed in the Model Section, the regression results confirm that the spatial model provides a better estimation of other socio-economic variables when the network effect is present. Table 2 shows that by controlling for spatial dependence, the estimates of the remaining explanatory variables have been changed (in most cases, the effects of other variables have been weakened). In addition, after accounting for spatial effects, a significant negative effect of education (variable "High Skilled") has been observed; beforehand, it was not significant, and it was weaker.

Thirdly, the spatial models provide a better data fit than do the traditional models. Since spatial models include an extra variable – weighted ethnic spatial lag – it inevitably leads to a likelihood gain. We compared Akaike's Information Criterion (AIC) results¹⁵ and likelihood ratio tests¹⁶ to investigate whether or not this gain is sufficient to overcome the penalty of the loss of the degree of freedom. In our analyses, spatial models in every case generate a lower AIC, which means the spatial model is the preferred method to model immigrants' self-employment decisions. Furthermore,

we followed Adjemian et al.'s (2010) approach to conduct a likelihood test on a vector of constraints equating the spatial models to the conventional model. Test results (including AIC measures) in column (3) and (4) of Table 2, compared to columns (1) and (2) for example, show that the spatial models are significantly different, and they can improve the likelihood of observing the original data as constraints were rejected ($\text{Prob} > \chi^2 = 0.00$) in all cases.

Furthermore, we report additional estimation results based on the assumption of exogeneity of the weighted ethnic spatial lag variable (Additional file 1: Table S2), supplementing the main results in columns (3) and (4) of Tables 2 and 3. Although both our exogenous and endogenous spatial models show a consistent pattern of the effects of ethnic network and other socio-economic variables, treating the ethnic network as an endogenous variable out-performs the alternative exogenous specification, as: (1) The Durbin and Wu–Hausman tests of endogeneity suggest control for endogeneity of ethnic network; and (2) The endogenous network effect spatial model always generated lower AIC results than both the exogenous network effect spatial model and also the conventional models. Therefore, our results are consistent with treating the spatial spill-over as multidirectional (rather than only unidirectional), and as such a much stronger network effect could be captured.

As discussed earlier, experience is argued to be either a “stock” (Evans 1989) or “flow” (Kidd 1993). In this study, we found that labour market experience works as a “stock” for recent immigrants in New Zealand. Labour market experience significantly increases the propensity for immigrants to be self-employed. With one more year’s labour market experience, the probability of self-employment will be increased by 0.15%. Therefore, with greater experience, immigrants tend to have obtained more management skills and better market information, which leads to greater skills for operating their own businesses.

The LisNZ data allow us to examine the effects of overseas self-employment experience on immigrants’ self-employment decisions in the host country. Again, the regression results suggest that one more year of overseas self-employment experience results in an increase in the probability of self-employment by 0.1%. Immigrants with experience of self-employment in their home country are more likely to be self-employed than are other immigrants without such experience.

Individuals gain knowledge of management and investment from education as well as from experience. On the one hand, educational attainment reflects learning, especially managerial skills to operate a business. On the other hand, individuals with higher educational attainment are less likely to be self-employed because higher educational attainment enhances the propensity for employment in the high-skilled waged sector. The latter hypothesis is confirmed by the empirical findings in this study. Keeping other factors constant, an immigrant with a high education level is less likely to be self-employed than an immigrant with lower educational attainments.

The coefficients of “years since migration” (YSM) in all models are positive at the highest level of statistical significance. By having one more years of living experience in the host country, immigrants can gain more information about the host labour market and business opportunities and can also expand their network. We find that the probability of self-employment is increased by 0.1% by one year’s increase in YSM. The positive effect of YSM observed in this panel study also confirmed the positive effect of YSM found by previous international studies (e.g. Li 2010; Toussaint-Comeau 2008).

In summary, as discussed before, immigrants may face disadvantages, such as lack of investment funds, information about the local market and its regulation, and less proficiency in the local language. The ethnic network can offset those disadvantages for immigrants to some extent. We found that, generally, the ethnic network significantly promotes self-employment for immigrants in New Zealand. Recent immigrants tend to share resources (e.g. financial resources, “protected market”, business opportunities, and local market information) through their own ethnic network and tend to help future immigrant entrepreneurs to become self-employed. Therefore, the self-employment decisions of recent immigrants are positively correlated. The ethnic network has a stronger positive effect for immigrants from traditional countries of origin than for non-traditional immigrants (immigrants from other regions except the United Kingdom).

5.2 By country of origin

The effects of ethnic capital may vary due to the culture, norms and other characteristics of different ethnic groups. In addition, immigrants from countries with language or institutional characteristics similar to the host country may show greater responsiveness to ethnic capital. For example, Chand and Ghorbani (2011) argued that the role played by social capital in the Chinese culture contrasts with that played in the Indian culture. They further show that Chinese immigrant entrepreneurs have a greater tendency than their Indian equivalents to take on employees from their ethnic community, which means that Chinese-operated businesses offer relatively greater paid job opportunities for Chinese immigrants.

Based on the endogenous network effect spatial model (as reported in column (4) of Tables 2 and 3), we expand our analysis to sub-samples by country of origin to examine whether we find differences across our ethnic groups. In particular, we examine the impact of the ethnic capital network variable in addition to the traditional ethnic concentration indicator and language proficiency on the likelihood of self-employment. For this we estimated the model for sub-samples of specified countries of origin of interest (UK and Ireland, China and India), and we compare results across these groups.

Results in Table 4 show that with one more year’s labour market experience, the probability of being self-employed increases much more for Chinese immigrants than it does for immigrants from other ethnic groups. The odds ratios for these results are presented in Additional file 1: Table S3. Likewise, Chinese immigrants are more likely to benefit from the experience of living in the host country than are other immigrants. Previous studies observed both the positive and negative effects of language skills on immigrants’ self-employment decisions. In this study, we found that proficiency in English has an insignificant negative effect on the self-employment decisions of recent male immigrants from China, while it positively influences the self-employment decisions of Indian immigrants. This shows that Indian immigrants with better language skills are more likely to be self-employed than are immigrants from China.

The effect of ethnic network is significant, positive and strong for immigrants in all cases. We further find that, compared to immigrants from non-English-speaking origins, those from the United Kingdom and Ireland and India (both members of the Commonwealth), Chinese entrepreneurs, in contrast derive greater benefit from ethnic concentration in a specific region. This result confirms the expectation that Chinese

Table 4 Self-employment decision by country of origin with network effect (average marginal effects)

	United Kingdom & Ireland	China	India
Ethnic capital			
Network effect†	0.03*** (0.011)	0.09*** (0.033)	0.04*** (0.013)
Ethnic concentration effect	-0.0007 (0.001)	0.003** (0.001)	-0.0008*** (0.0003)
Selected human capital			
Experience	0.0016*** (0.0002)	0.0024*** (0.0007)	0.0004*** (0.0001)
Self-employment experience (origin)	0.0009*** (0.0002)	0.0008** (0.0004)	0.0009*** (0.0003)
High skilled	0.0005 (0.0007)	-0.054*** (0.0018)	-0.002** (0.0009)
English proficiency	/	-0.0002 (0.002)	0.0013*** (0.0004)
Manager & professional	0.0006 (0.0005)	0.0057*** (0.002)	0.0004 (0.0004)
Selected personal characteristics			
Years Since Migration (YSM)	0.0008*** (0.0002)	0.0040*** (0.001)	0.0004 (0.0003)
Married	0.0032*** (0.0006)	0.0021** (0.001)	0.0010** (0.0004)
Children	-0.0031*** (0.0006)	-0.004*** (0.0014)	-0.002*** (0.0006)
Own dwelling	0.0015** (0.0006)	-0.0034*** (0.001)	0.0001 (0.0001)

Note: †Network (weighted ethnic spatial lag 'Wy') is treated as an endogenous variable.

(1) Panel Logit estimations.

(2) Standard errors in parentheses.

(3) *p<0.10 **p<0.05 ***p<0.01.

immigrants are more likely to benefit from a larger ethnic enclave in this setting. With an increase in the ethnic population, the ethnic enclave can generate further business opportunities for Chinese immigrants.

Overall, the combination of the ethnic Network Effect, and the Ethnic Concentration Effect variables confirms that these factors increase the propensity of immigrant self-employment through alternative means. Our approach that allowed accounting for both sets of variables allowed us to observe their separate effects.

6. Conclusion

The thrust of this paper is that while immigrants face many barriers to employment in the host country's labour market, ethnic capital (e.g. network and ethnic concentration) may help immigrants to overcome those disadvantages to some extent. By choosing a specific location, immigrants can benefit from ethnic capital and face fewer barriers to employment (e.g. self-employment).

Previous international studies (e.g. Bertrand et al. 2000; Edin et al. 2003) adopt either ethnic concentration or language as a proxy for the immigrants' network in the host country. We have provided an alternative approach for measuring how the self-employment decision made by ethnic-neighbours influences a typical immigrant's self-employment choices. We contribute to this literature by adopting the "spatial model approach" to account for ethnic concentration and networks in order to capture the effects of social and resource networks for immigrant groups. We incorporate different measures of ethnic capital, in particular ethnic group economic resources and spatial-ethnic concentration. In this paper, we compare Akaike's Information Criterion (AIC) results and likelihood ratio tests to investigate whether or not the gain from having the "weighted ethnic spatial lag" (rather than having the conventional network variable) is sufficient to overcome the penalty of the loss of the degree of freedom. In our analyses, spatial models, in every case, generate a lower AIC, which means the spatial model is the preferred method for modelling immigrants' self-employment decisions. Furthermore, we follow Adjemian et al.'s (2010) approach to conduct a likelihood test on a vector of constraints, comparing the spatial models to the conventional model. We show that the spatial models are significantly different and that they considerably improve the likelihood of observing the original data.

As a result, compared to the conventional approach, the spatial approach we used in this paper offers three advantages: it captures a much more accurate effect of the network resources; it provides a better estimation of the effect of other socio-economic variables; and it provides a better data fit. Furthermore, the weighted ethnic spatial lag variable we recommend includes consideration of other immigrants' choices and location, not just neighbourhood.

The empirical findings of this study strongly suggest that an ethnic network promotes self-employment among recent immigrants. It implies immigrants may share socio-economic resources through their ethnic network and may help one another for self-employment. The extent of this impact varies by ethnic group. We find that the exclusion of the network correlation effect results in the underestimation of other important variables such as education and skill.

Endnotes

¹Immigrants are more likely to be economically connected with their own ethnic group in the same area (e.g. Battu et al. 2011).

²Metropolitan Statistical Area.

³We derive this variable from the published New Zealand Census tables. The New Zealand Census provides information for the entire resident New Zealand population. The number of people who usually reside in a locality is reported by gender and the country of origin in both the 1996 and 2001 published New Zealand census tables. Based on the Census, we are able to identify and incorporate 26 countries of origin. These include the key countries of interest as immigrant source countries in our analyses, such as the UK, Ireland, China, and India.

⁴Since we test the lagged value of ethnic concentration in this study, the lagged ethnic concentration variable is treated as an exogenous variable here.

⁵Recent studies have also controlled for the spatial lag variable as endogenous variable (e.g. Goetzke and Andrade 2009).

⁶We have conducted the Durbin and Wu–Hausman tests of endogeneity, and both tests suggest control for endogeneity of ethnic network.

⁷All immigrants (except refugees) more than 15 years old who were approved for permanent residence in New Zealand are included in the target population for LisNZ.

⁸The majority of American studies (e.g. Yuengert 1995) have examined immigrants' geographical decisions in light of Metropolitan Statistical Areas. We would argue that New Zealand's 'regions', as generally organised around a major city, are quite equivalent, in terms of size. Given that this paper emphasizes the impact of ethnicity and location on employment levels, it follows that we consider immigrants' self-employment decisions at the New Zealand regional level. As such, immigrants living in one suburb of a region are able to access knowledge/information, indirectly, regarding self-employment in other suburbs of that region, and a network is likely to develop within that region. For example, Immigrant A, located in a northern suburb has a friend, Immigrant B, located in the Central suburb, connected to a third friend, Immigrant C, residing in a southern suburb. Therefore, Immigrant A may obtain information about self-employment opportunities in the southern suburb through the network. Taking into account this information, and adopting our ethnic spatial weighted matrix, we aim to capture the impact of an entire network on a particular location. Another relevant scenario is that immigrants might cluster in one region, but be self-employed elsewhere. For example, some immigrants may live in one suburb, yet work in another suburb. Therefore, we think in considering their locations, the regional perspective helps to best address this in this setting.

⁹According to the 2006 New Zealand Census (Statistics New Zealand, 2006), China has become the second largest immigrant source country for New Zealand (after the United Kingdom and Ireland). In our analyses we consider the major ethnic groups from the UK and Ireland, China, India, the Pacific Islands, and the rest of the world.

¹⁰In order to investigate the robustness of the spatial model results, we also estimate the spatial model on the relevant sample which also includes unemployed immigrants whose ages are between 20 and 55 years old. The regression results confirm the significant positive effect of ethnic network on recent immigrants' self-employment decision.

¹¹As noted earlier, the first survey was conducted six months after new immigrants settled in New Zealand. Therefore, binary variables such as English proficiency, children, marriage, high skilled, own dwelling and managerial and professional previous occupation in our model are treated as exogenous variables, as based on only the first wave's data.

¹²We adopt the random-effects setting, the variation across immigrants is treated as random, and the unobserved effect is also assumed to be uncorrelated with the explanatory variables.

¹³Our selection of estimation method was based on a full consideration of alternative estimation methods. For example, the Fixed-effects method commonly used in panel data settings is not suitable in this case since besides the effects of ethnic capital, it is

important to control for the impact of other human capital variables on immigrants' self-employment decisions. In this setting, to avoid endogeneity, the measures of English proficiency, high-skilled, children, marriage, assets (e.g. own a property) and previous occupation are based on the initial wave interviews, and as such they are time-invariant, making Fixed-effects estimation inappropriate.

¹⁴We have also performed a robustness check and tested the Toussaint-Comeau's (2008) approach to measuring the quality of network by a linear variable based on the difference between the average self-employment rate of the ethnic group and the general average self-employment rate in the host country. Regression results suggest a significant positive effect of network, which confirms the finding from the spatial models. However, the spatial approach we have adopted is still preferred as it generates lower Akaike's Information Criterion (AIC) results in all cases examined.

¹⁵The AIC, or Akaike Information Criterion, provides a way of measuring a statistical model, in terms of its relative quality, for a specific collection of data. As such, it enables the selection of models. It does not allow for the testing of a model, in terms of investigating a hypothesis. However, it is appropriate when the elements of usefulness/appropriateness versus complexity are taken into consideration. The $AIC = 2k - 2\ln(L)$, where k represents the model's number of parameters, and L represents the ultimate value of the function of the model's likelihood. Adjemian et al. (2010) adopt this approach to select the best model for an individual's choice of automobile, when considering conventional and spatial models.

¹⁶A likelihood ratio test employs statistics to compare the fit of two different models, where one (the null model) represents a particular variation of the other (the alternative model). The test employs the ratio of likelihood, comparing the occurrence of data across the two models.

Additional file

Additional file 1: Table S1. Descriptive statistics of employed males. **Table S2.** Estimates of self-employment decision with network effect. **Table S3.** Self-employment decision by country of origin with network effect (odds ratios).

Competing interests

The IZA Journal of Migration is committed to the IZA Guiding Principles of Research Integrity. The authors declare that they have observed these principles.

Acknowledgement

We would like to express appreciation to Professors Daniel Hamermesh, Deborah Cobb-Clark, John Haisken-DeNew and fellow researchers at the Melbourne Institute of Applied Economic and Social Research for providing feedback regarding our research. Thanks must also go to the New Zealand Ministry of Business, Innovation and Employment Labour Group and Statistics New Zealand, who provided access to the data used in this paper.

Access to the data used in this study was provided by Statistics New Zealand under conditions designed to give effect to the security and confidentiality provisions of the Statistics Act 1975. All non-regression results using Longitudinal Immigrant Survey: New Zealand (LisNZ) data are subject to base three rounding in accordance with Statistics New Zealand's release policy for LisNZ data. Any views expressed are the sole responsibility of the authors and do not purport to represent those of Ministry of Business, Innovation and Employment, or Statistics New Zealand.

Responsible editor: Denis Fougère

Author details

¹Department of Economics, The University of Auckland, 12 Grafton Road, Auckland 1010, New Zealand. ²Graduate School of Management, The University of Auckland, 12 Grafton Road, Auckland 1010, New Zealand.

Received: 3 April 2014 Accepted: 25 August 2014

Published online: 01 October 2014

References

- Adjemian MK, Lin CYC, Williams J (2010) Estimating spatial interdependence in automobile type choice with survey data. *Transport Res A-Pol* 44:661–675
- Aguilera MB (2009) Ethnic enclaves and the earnings of self-employed Latinos. *Small Bus Econ* 33:413–425
- Aldrich H, Cater J, Jones T, McEvoy D, Velleman P (1985) Ethnic residential concentration and the protected market hypothesis. *Soc Forces* 63(4):996–1009
- Aldrich H, Jones T, McEvoy D (1984) Ethnic advantage and minority business development. In: Ward R, Jenkins R (eds) *Ethnic Communities in Business*. Cambridge University Press, New York, pp 189–210
- Aldrich H, Waldinger RH (1990) Ethnicity and entrepreneurship. *Annu Rev Sociol* 16:111–135
- Andersson L, Hammarstedt M (2011) Ethnic enclaves, networks and self-employment among Middle Eastern immigrants in Sweden. *Int Migr*. doi:10.1111/j.1468-2435.2011.00714.x
- Anselin L (1990) Some robust approaches to testing and estimation in spatial econometrics. *Reg Sci Urban Econ* 20:141–163
- Anselin L (1999) *Spatial Econometrics*. University of Texas at Dallas, Bruton Center, School of Social Sciences
- Anselin L (2002) *Spatial regression analysis short course*. Unpublished manuscript, University of Michigan, Ann Arbor
- Anthony DL (1999) The importance of trust in micro-credit borrowing groups. Paper presented at the Federal Reserve System Conference on Business Access to Capital and Credit, Arlington VA
- Battu H, Seaman P, Zenou Y (2011) Job contact networks and the ethnic minorities. *Labour Econ* 18(1):48–56
- Bernhardt I (1994) Comparative advantage in self-employment and paid work. *Can J Econ* 12(4):273–289
- Bertrand M, Luttmer E, Mullainathan S (2000) Network effects and welfare cultures. *Q J Econ* 115(3):1019–1055
- Bonacich E, Modell J (1980) The economic basis of ethnic solidarity: Small business in the Japanese American community. University of California Press, Berkeley
- Bond P, Townsend R (1996) Formal and informal financing in Chicago neighborhood. *Econ Perspect* 20(4):3–27
- Borjas GJ (1986) The self-employment experience of immigrants. *J Hum Resour* 21(4):485–506
- Borjas GJ (1992) Ethnic capital and intergenerational mobility. *Q J Econ* 107(1):123–150
- Borjas GJ (1995) Ethnicity, neighborhoods, and human capital externalities. *Am Econ Rev* 85(3):365–390
- Boyd RL (1990) Black business transformation, black well-being, and public policy. *Popul Res Policy Rev* 9:117–132
- Bradley DE (2004) A Second Look at Self-Employment and the Earnings of Immigrants. *Int Migr Rev* 38(2):547–583
- Chand M, Ghorbani M (2011) National culture, networks and ethnic entrepreneurship: A comparison of the Indian and Chinese immigrants in the US. *Int Bus Rev* 20(6):593–606
- Chiswick BR (1978) The effect of Americanisation on the earnings of foreign-born men. *J Polit Econ* 86(5):897–921
- Clark K, Drinkwater S (1998) Ethnicity and Self-Employment in Britain. *Oxford B Econ Stat* 60:383–407
- Clark K, Drinkwater S (2000) Pushed out or pulled in? Self-employment among ethnic minorities in England and Wales. *Labour Econ* 7(5):603–628
- Damm A (2009) Determinants of recent immigrants' location choices: quasi-experimental evidence. *J Popul Econ* 22(1):145–174
- Edin PA, Fredriksson P, Aslund O (2003) Ethnic enclaves and the economic success of immigrants-Evidence from a natural experiment. *Q J Econ* 118:329–357
- Evans MDR (1989) Immigrant entrepreneurship: Effects of ethnic market size and isolated labour pool. *Am Sociol Rev* 54(6):950–962
- Goetzke F (2008) Network effects in public transit use: Evidence from a spatially autoregressive mode choice model for New York. *Urban Stud* 45(2):407–417
- Goetzke F, Andrade PM (2009) *Progress in Spatial Analysis*. In: Snickars F, Thill JC, van Dijk J, Westlund H (eds) *Fischer MM. Advances in Spatial Science*, Springer, pp 217–229
- Jean S, Jimenez M (2011) The unemployment impact of immigration in OECD countries. *Eur J Polit Econ* 27(2):241–256
- Kelejian H, Prucha I (1998) A generalized spatial two stage least squares procedure for estimating a spatial autoregressive model with autoregressive disturbances. *J Real Estate Financ* 17:99–121
- Kelejian H, Robinson DP (1993) A suggested method of estimation for spatial interdependent models with autocorrelated errors, and an application to a county expenditure model. *Pap Reg Sci* 72:297–312
- Kidd MP (1993) Immigrant wage differentials and the role of self-employment in Australia. *Aust Econ Pap* 32(60):92–115
- Le AT (2000) The determinants of immigrant self-employment in Australia. *Int Migr Rev* 34(1):183–214
- LeSage J, Pace RK (2009) *Introduction to Spatial Econometrics (Statistics: A Series of Textbooks and Monographs)*. Chapman & Hall/CRC, 6000 Broken Sound Parkway NW, Suit 300, Boca Raton, FL 33487-2724, US
- Li PS (2010) Immigrants' propensity to self-employment: Evidence from Canada. *Int Migr Rev* 35(4):1106–1128
- Light I (1979) Disadvantage minorities in self-employment. *Int J Comp Sociol* 20(1–2):31–45
- Light I, Sanchez AA (1987) Immigrants entrepreneurs in 272 SMSAs. *Sociol Perspect* 30:373–399
- Lofstrom M (2002) Labor market assimilation and the self-Employment decision of immigrant entrepreneurs. *J Popul Econ* 15(1):83–114
- Maré D, Stillman S (2009) The labour market adjustment of immigrants in New Zealand. *Economic Impacts of Immigration Working Paper Series*. Motu Economic and Public Policy Research, Wellington
- Nana G, Sanderson K (2008) *Migrants and labour market outcomes. Economic Impacts of Immigration Working Paper Series*. New Zealand Department of Labour, Wellington, New Zealand
- Organisation for Economic Co-Operation and Development (2001) *The employment of foreigners: Outlook and issues in OECD countries*. Retrieved from www.oecd.org/employment/emp/2079451.pdf
- Poot J (1998) The impact of immigration on labour markets and urban infrastructure in Australia and New Zealand. In: Gorter C, Nijkamp P, Poot J (eds) *Crossing borders: Regional and urban perspectives on international migration*. Ashgate, Aldershot UK, pp 129–157
- Raijman R, Tienda M (2003) Ethnic foundations of economic transactions: Mexican and Korean immigrant entrepreneurs in Chicago. *Ethnic Racial Stud* 26(5):783–801

- Sousa LD (2013) Community determinants of immigrant self-employment: Human capital spillovers and ethnic enclaves. US Census Bureau Center for Economic Studies Paper No. CES-WP-13-21
- Toussaint-Comeau M (2008) Do ethnic enclaves and networks promote immigrant self-employment? *Econ Perspect* 32(4th Quarter):30–50
- Van Auken HE, Neeley L (1998) Evidence of bootstrap financing among small start-up firms. *J Entrep Small Bus Financ* 5(3):235–249
- Waldinger RH (1986) Immigrant enterprise, a critique and reformulation. *Theor Soc* 15(1–2):249–285
- Wilson K, Portes A (1980) Immigrant enclaves: An analysis of the labor market experiences of Cubans in Miami. *Am J Sociol* 86:195–319
- Yuengert AM (1995) Testing hypotheses of immigrant self-employment. *J Hum Resour* 30(1):194–204

doi:10.1186/s40176-014-0018-3

Cite this article as: Wang and Maani: **Ethnic capital and self-employment: a spatially autoregressive network approach.** *IZA Journal of Migration* 2014 **3**:18.

Submit your manuscript to a SpringerOpen[®] journal and benefit from:

- ▶ Convenient online submission
- ▶ Rigorous peer review
- ▶ Immediate publication on acceptance
- ▶ Open access: articles freely available online
- ▶ High visibility within the field
- ▶ Retaining the copyright to your article

Submit your next manuscript at ▶ springeropen.com
