I Z A Institute
of Labor Economics
Initiated by Deutsche Post Foundation

DISCUSSION PAPER SERIES

IZA DP No. 14862

Ethnic Regional Networks and Immigrants' Earnings: A Spatial Autoregressive Network Approach

Xingang Wang
Sholeh A. Maani

NOVEMBER 2021

I Z A Institute
of Labor Economics

Initiated by Deutsche Post Foundation

DISCUSSION PAPER SERIES

IZA DP No. 14862

Ethnic Regional Networks and Immigrants' Earnings: A Spatial Autoregressive Network Approach

Xingang Wang
University of Auckland

Sholeh A. Maani
University of Auckland and IZA

NOVEMBER 2021

IZA – Institute of Labor Economics

Schaumburg-Lippe-Straße 5–9
53113 Bonn, Germany

Phone: +49-228-3894-0
Email: publications@iza.org

www.iza.org

ABSTRACT

Ethnic Regional Networks and Immigrants' Earnings: A Spatial Autoregressive Network Approach

The conventional model of immigrant earnings does not account for the correlation of outcomes across immigrant ethnic networks. We apply a spatial autoregressive network approach to account for the spill-over effects of migrant ethnic group economic resources and labour market outcomes. We employ unit-record data across 10 years for New Zealand, a major immigrant receiving country. By applying generalised method of moment (GMM) estimation, we address endogeneity of the spatial network variable. Results confirm strong positive associations of earnings with both ethnic concentration and networks of resources. The analytically enhanced approach provides opportunities for new research on the determinants of immigrant earnings.

JEL Classification: J30, J31, Z13, Z18

Keywords: immigrant, earnings, ethnic network, spatial autoregressive model, GMM estimation

Corresponding author:
Sholeh A. Maani
Department of Economics
University of Auckland
12 Grafton Road
Auckland 1010
New Zealand
E-mail: s.maani@auckland.ac.nz

1 Introduction

Is settlement in locations with high ethnic concentration positively or negatively associated with immigrants' earnings? The emerging international evidence on this question is unresolved. As immigrants now compose a significant and increasing proportion of populations in Western countries, this question is worth close examination across immigrant-receiving countries. Indeed, how well immigrants perform in their host country's labour market is an important factor that influences their post-migration integration and economic success, as well as the economic productivity of their destination (host) country. [1]

In this paper, we consider ethnic resources and networks as key concepts. We enhance the conventional migrant economic performance model by incorporating a spatial autoregressive network approach (Adjemian et al. 2010; Goetzke 2008; Goetzke and Weinberger 2012; LeSage and Pace 2009) to incorporate the impact of immigrants' ethnic resources and networks on their labour market performance. An important feature of this approach is that it enables us to formally incorporate the co-dependence of economic outcomes for immigrant ethnic groups, based on the ethnic group's network of economic resources. The conventional econometric model of immigrant earnings does not account for the correlation of outcomes across immigrant networks.

It is generally accepted in the migration economics literature that new immigrants may face earnings disadvantages in their new country. For example, immigrants may have deficiencies in local language skills and social networks, lack information about job opportunities and local employers, and be less familiar with institutional factors and cultural customs (e.g., Borjas 1985; Chiswick 1978; Chiswick et al. 2005; De Jong and Madamba 2001; Dustmann and Fabbri 2003).

An increasing number of studies have paid attention to the differences of earnings across immigrant ethnic groups (e.g., Beenstock et al. 2010; Borjas 1985; Kossoudji 1988;

[1] 'Host' country, as in the migration economics and demography literature refers to a migrant's country of residence (host), as opposed to 'home' country, which refers to the migrant's country of birth or pre-migration home country.

McDonald and Worswick 1999). It is also recognised that the earnings growth of different ethnic groups has diverse patterns and time ranges. Borjas (1985) observed divergent assimilation processes for immigrants from Cuba and Mexico to the United States. McDonald and Worswick (1999) documented the persistence of income differences between immigrants from a non-English-speaking background and natives in Australia. Studies related to New Zealand have confirmed differences in immigrant earnings and assimilation effects across ethnic groups (e.g., Masgoret et al. 2012; Poot 1998; Stillman and Maré 2009).

These findings give rise to questions as to why there are differences in economic performance across ethnic groups and how ethnic group membership influences immigrants' labour market performance. Furthermore, previous studies have assumed that labour market performance data for individuals are independent and identically distributed (i.i.d.). However, one may consider whether individuals within spatial ethnic groups influence each other and if their labour market performance is correlated to some extent.

International evidence further confirms that immigrants tend to reside in a few localities or states in immigrant-receiving countries. This effect is significant in the United States (US), as originally noted by Borjas (1999), where the greatest majority of immigrants resided across only six states. Other countries show similar effects (e.g., Dustmann and Fabbri 2003 for the United Kingdom (UK), and Edin et al. 2003 for Sweden). New Zealand, as an English-speaking and immigrant-receiving country, is very similar to other immigrant-receiving countries in this respect (e.g., Masgoret et al. 2012). Therefore, ethnic concentration is a component of the economic experiences of, and outcomes for, immigrants in their destination country.

Studies that have investigated the impact of ethnic concentration find varied results on whether residence in areas with strong linguistic or ethnic concentration is beneficial or harmful to immigrants' economic success (e.g., Beckhusen et al. 2013; Bertrand et al. 2000; Chiswick and Miller 2002a; Cuttler and Glaeser 1997; Edin et al. 2003; Warman 2007; Zhu et al. 2013).[2] There are competing theoretical arguments for the effect of ethnic concentration on economic outcomes. Ethnic concentration can provide significant new

[2] This is not intended as an exhaustive representation of related studies.

market opportunities (Portes 1987; Portes and Shafer 2007). However, factors such as competition due to over-supply, lack of integration into the mainstream economy and slow language proficiency development are among the reasons why the overall impact may be negative (Beckhusen et al. 2013; Bertrand et al. 2000; Chiswick and Miller 1996, 2002b). Edin et al. (2003) find positive effects for the higher end of the earnings distribution.[3]

An important factor that also influences these outcomes for immigrants, and is receiving recent attention, is the strength of economic resources and social networks that they can access within their ethnic network. Recent studies show that social networks can exert a significant positive influence on individuals' labour market performance (Baltagi et al. 2017; Battu et al. 2011; Edin et al. 2003; Piracha, et al. 2016).

In addition, recent studies have increasingly noted the need for analyses that recognise the potential connection of individuals beyond the size or density of the network. For example, individuals may benefit from their friendships, such as their friends may introduce job opportunities to them or assist them with their job applications. Social networks are argued to be the most profitable avenue of job search for immigrants (Frijters et al. 2005). Compared to the native-born population, immigrants are possibly more dependent on their social networks for greater economic integration in the host country, because they usually have less knowledge of the host country's labour market. For these reasons, the labour market performances of individual immigrants are expected to be correlated with each other. In particular, ethnic social networks may have a positive impact on the process of immigrants' assimilation. Other strands of research in sociology and population studies also note the potential economic impacts of ethnic networks for immigrants (e.g., Fong and Jing 2011; Portes 1987; Portes and Shafer 2007).[4]

This effect is distinct from modelling with conventional ethnic concentration ratios or related indicators, such as size of the network, information, and job referrals.

[3] Their study used a natural experiment where the identification of the system relied on refugees who had stayed in their assigned locality.
[4] These examples are representative; they are not intended to be exhaustive.

In a different strand of research, studies that formally incorporate spatial economic factors have received recent attention across economic and social dimensions (e.g., Adjemian et al. 2010; Goetzke 2008; Goetzke and Rave 2015; Goetzke and Weinberger 2012; Kalenkoski and Lacombe 2013). This spatial autoregressive network approach, which incorporates a weight matrix, has made recent contributions in transportation economics by formally incorporating the impact of economic networks into economic models of interest. A major added advantage of this approach is that it lends itself to addressing the endogeneity of spatial social networks through generalised method of moment (GMM) estimation. The application of this spatial approach to labour market outcomes provides new opportunities that are particularly relevant for research on immigrant labour market performance. We apply the spatial approach to enhance the conventional models of immigrant earnings. In considering the effect of ethnic networks in labour market settings, accounting for the special nature of data is non-trivial. In this setting, labour market outcomes are expected to be influenced by spill-over effects that are geographically relevant to an individual. Notably, the spatial aspect of immigrants' labour market can play a relevant role in relation to information and job opportunities that are geographically viable for immigrants.

The contributions of this paper to the literature are as follows: Conceptually, the analysis expands the conventional economic analysis of immigrant earnings to incorporate the impact of spatial ethnic economic networks. Empirically, the spatial weight matrix of ethnic networks adopted in this study controls for the potential co-dependence of immigrant labour market outcomes, and it reflects differences of group resources across ethnic groups and locality. Importantly, the analysis also accounts for endogeneity of the network variable. In our analysis, we consider the association of both ethnic concentration and the ethnic-spatial economic network with earnings, based on country of origin. The paper further contributes to the literature by providing new evidence on the effect of ethnic network outcomes by country-of-origin language group for immigrants residing within and outside of their ethnic group's high ethnic concentration areas.

To the best of our knowledge, this is the first application of the spatial autoregressive network approach to immigrant earnings.

We use individual-level data from a rich New Zealand data set, the New Zealand Income Survey (NZIS). We use data across 10 years (2001–2010), matched with relevant regional-level ethnic concentration data from the 1996 and 2001 censuses. Much of the earlier analyses of immigrant economic performance have been based on US and UK data, and expanding the analysis to other major immigrant-receiving countries is desirable. New Zealand is a major English-speaking immigrant-receiving country, with a long history of migration. A positive feature of using New Zealand data for this analysis is that New Zealand receives large proportions of immigrants from diverse linguistic and ethnic backgrounds. In particular, 36 countries of birth reported for immigrants include three major groups of immigrants with an extended history of migration – from the UK and Ireland; the Pacific Islands; and Asia – providing a diversity of ethnic group and language networks, ethnic-spatial concentrations, and labour market evidence for this analysis.

Our results confirm that the adopted spatial autoregressive ethnic network approach offers two advantages: It allows researchers to estimate the potential economic spill-over effects of the immigrants' ethnic group network; and compared to the conventional approach, the spatial approach provides a better data fit. Notably, we find positive associations between both ethnic concentration and the network of economic resources and immigrant earnings. For immigrants from a non-English-speaking background, and Asian immigrants in particular, living in an enclave region is associated with significant and high economic spill-over effects. Since New Zealand has skilled migration selection policies, our results suggest that positive effects to the economy from both ethnic concentration and the strength of immigrant economic resources can supersede potential negative effects.

This paper is arranged as follows. After the introduction, Section Two provides a brief description of relevant hypotheses based on that concept. Section Three discusses the spatial econometric models. Section Four describes the data set, followed by results in Section Five. Section Six provides concluding remarks.

2 Immigrant ethnic concentration and economic networks

Ethnic concentration and ethnic-spatial networks of economic resources measure different aspects of the geographic concentration effect. Ethnic concentration is a measure of the relative size of an immigrant ethnic population, relative to the overall population in a geographic area. As such, it may be considered as a measure of the probability that an immigrant would be able to meet or benefit from resources belonging to the network. This measure has been considered in earlier studies as a measure of enclave effects.

While ethnic concentration may be considered as a measure of the relative size of an immigrant ethnic group in a geographic area, ethnic group resources generally represent the strength of economic resources shared by an immigrant ethnic network in a geographic area.

Below we briefly discuss the relevance of both ethnic concentration and ethnic spatial network of economic resources as separate measures.

Ethnic concentration

The effects of ethnic concentration on immigrants' earnings might be different across ethnic groups. If immigrants compete for scarce employment opportunities in the host country's labour market, immigrants may accept a lower wage. Similarly, if an increase in the proportion of immigrants in a specific region generates higher demand for immigrant labour and greater job opportunities, a higher wage for immigrants is expected.

Therefore, the impact of ethnic concentration on immigrants' economic performance continues to be an empirical question, and it may vary by host country and ethnic group. In particular, if the quality (strength) of ethnic group resources and the co-dependence of outcomes are not controlled for, ethnic concentration alone may provide mixed results by combining ethnic concentration and ethnic resource effects. [5]

[5] The term 'quality' as in the migration economics literature (going back to the work of Borjas (1985)), refers to greater resources of immigrants.

Ethnic economic networks

The importance of ethnic group economic resources in the economics literature was emphasised by Borjas (1992, 1995) in immigrant economic assimilation/integration studies, and he referred to these economic resources as "ethnic capital". The theory refers to a type of social capital and resources belonging to a certain ethnic group. Borjas (1992) attributed a main reason for different earnings and economic assimilation profiles across ethnic migrant groups to this effect. It includes factors (e.g., the level of economic development, native language, and political system) that stem from the country of origin and act as an "innate" capital (and resource) for immigrants who originate from their source country. Ethnic capital is a resource, as well as a capital, accessible by immigrants from the same group.

The spatial concentration of ethnic resources was not a major feature of the earlier ethnic capital analyses. For example, Borjas' (1992) seminal analysis of ethnic capital employed national-level average earnings by ethnic group.

Other strands of research (e.g., Bonacich and Modell 1980; Chiswick and Miller 2002a; Fong and Jing 2011; Portes 1987; Portes and Shafer 2007) emphasise the social bonds between immigrants from the same cultural or linguistic background. According to the ethnic network hypotheses put forward by Munshi (2003), Piore (1979), and Portes and Shafer (2007) , ethnic networks assist immigrants to adjust to the new environment in the host country. They enhance new immigrants' feelings of security and identity within the ethnic group, since they share the same culture, language, and norms with each other. At the initial stage of immigration, an ethnic network plays a vital role in helping a new immigrant settle down and access employment opportunities. Recent studies show positive evidence on the role of social networks on an individual's labour market performance through finding employment (e.g., Frijters et al. 2005).

We acknowledge that ethnicity may represent different meanings to different people, based on race, language, country of origin, etc. We also acknowledge that not everyone born in the same country may have the same racial or language background. We are using the term ethnic concentration to refer to the positive links and a sense of

belonging that comes with a shared country of origin. Using the criterion of country of birth, as we do in this study, has advantages in our study's setting. Notably, from an econometric point of view, country of birth (as opposed to self-reported ethnicity) is advantageous for exogeneity criteria.

The present paper enhances the economic definition of ethnic capital, extending the concept to include a network of resources based on shared beliefs, group language proficiency and skills, and other economic resources available to a typical migrant ethnic group, with geographic concentration. This resource, generated from the ethnic environment in the host country, is influenced by local socio-economic factors. In this setting, the spatial aspect of immigrant resources is also an important component of the economic network. In addition, this resource comes from immigrants themselves, so it therefore implies that the "innate" capital from the host country may vary over time, which is different from the nature of the "innate" capital from the country of origin.

In our modelling approach to immigrant earnings, we account for this indicator, representing the ethnic-spatial network of economic resources, in addition to ethnic concentration.

3 Model and estimation method

The conventional form of the model to analyse how immigrant earnings respond to the assimilation/integration processes in the economics literature is:

$$Y_i = \beta_1 X_i + \gamma YSM_i + \beta_2 M_i + \varepsilon_i \quad (1)$$

where Y_i denotes the natural logarithm of hourly earnings of individual i in the host country; X is a vector of explanatory variables (e.g., years of schooling, labour market experience); YSM denotes years since migration to the host country; M is a dummy variable, set to 1 if foreign-born and 0 otherwise; ε_i is the random error term; and γ measures how earnings grow with the integration process (Chiswick 1978; Borjas 1985).

In this paper we augment the conventional econometric model of immigrant earnings to incorporate two different measures of spatial economic connection: immigrant country-of-birth group economic resources and ethnic concentration

3.1 Ethnic-spatial approach

Previous international studies adopted either ethnic concentration (e.g., Logan et al. 2003) or language (e.g., Chiswick and Miller 1996, 2002b) as a proxy of the immigrant's network in the host country.

We incorporate the spatial weight matrix approach to account for ethnic concentration and networks in order to capture the effects of social and resource networks for immigrant groups. The spatial lag matrix is constructed with microdata and is based on the three conditions of ethnic group, region of residence, and year of survey. As such, the matrix has different values by country-of-birth group and geographic location, and it reflects changes in the resources of the network matrix over time. As discussed in Section Two, we use country-of-birth group as a proxy for ethnic groups residing in the same area.

Under the ethnic-spatial network hypothesis, an individual's earnings depend on their ethnic group resources 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-spatial group members. Thus, "ethnic-spatial dependence" represents the case that an individual's labour market performance is influenced by their ethnic-spatial group members' labour market performance, reflecting the group's economic resources network (ethnic capital) in that location.

We set W as an $n \times n$ ethnic-spatial weight matrix that shows the first-order ethnic and geographical (ethnic-spatial) relationship among individuals. Before the discussion of W, the first-order ethnic-spatial member matrix E is relevant. Suppose P1, P2, P4, and P6 are all from the UK, P1 and P4 are located in region A, and P2 and P6 are located in region B. Suppose P3, P5, and P7 are from Asia and all of them are located in region B. Thus, the 7×7 first-order ethnic-spatial group member matrix E is:

$$E = \begin{array}{c} \\ P1 \\ P2 \\ P3 \\ P4 \\ P5 \\ P6 \\ P7 \end{array} \begin{array}{c} \begin{array}{ccccccc} P1 & P2 & P3 & P4 & P5 & P6 & P7 \end{array} \\ \begin{pmatrix} 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 1 \\ 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 1 \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 1 & 0 & 0 \end{pmatrix} \end{array} \tag{2}$$

When the elements of the matrix *E* are zeros, individuals are deemed not to be the first-order ethnic-spatial group members (neighbours in a broad sense). In addition, the diagonal elements of the above matrix are zeros, such that individuals are not considered as spatial group members to themselves.

In order to define an "ethnic-spatial lag", the ethnic-spatial matrix *E* should be normalised by unifying the row sums and so we can form the ethnic-spatial weight matrix *W*:

$$W = \begin{array}{c} \\ P1 \\ P2 \\ P3 \\ P4 \\ P5 \\ P6 \\ P7 \end{array} \begin{array}{c} \begin{array}{ccccccc} P1 & P2 & P3 & P4 & P5 & P6 & P7 \end{array} \\ \begin{pmatrix} 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1/2 & 0 & 1/2 \\ 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1/2 & 0 & 0 & 0 & 1/2 \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1/2 & 0 & 1/2 & 0 & 0 \end{pmatrix} \end{array} \tag{3}$$

The W matrix is the ethnic-spatial lag operator, which when applied on a specific vector y, for instance, produces the ethnic-spatial lag of this vector.

In our analysis, the *E* and W matrices are specified based on 36 distinct country-of-birth groups of immigrants, 16 regions for each year, over ten years of data. Therefore, different country-of-birth groups have different weights assigned to them in the matrices by country-of-birth group, region and time. As a result, the weights in the *W* matrix for immigrants from different country-of-birth group in the same location and year are different. This formulation, that incorporates different weights in the W matrix for different groups of immigrants is more precise for controlling for group heterogeneity,

and consistent with spatial approaches in Anselin (1990), and the geographically weighted regression approach (Fotheringham, Brundson and Charlton, 2002).

LeSage and Pace (2009) indicate the data-generating process when the value of an observation *i* depends on the value of its neighbour *j*'s observation:

$$y_i = \alpha_i y_j + \beta X_i + \varepsilon_i \quad (4)$$

$$y_j = \alpha_j y_i + \beta X_j + \varepsilon_j \quad (5)$$

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

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

Thus, Equations 4 and 5 imply a "simultaneous data generating process", showing the dependence of y_i and y_j and vice versa. This feature leads us to a data-generating process that is an "ethnic-spatial autoregressive process", and we can have the following expression:

$$y_i = \rho \sum_{j=1}^{n} W_{ij} y_j + \varepsilon_i \quad (6)$$

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

An "ethnic-spatial group member " is defined as two or more individuals who are from the same ethnic group and settled in the same location at a specific time; thus, $\sum_{j=1}^{n} W_{ij} y_j$ is the "ethnic-spatial lag" in this setting, and it represents the linear combination of individual *i*'s ethnic neighbours' labour market performances.

As a result, the matrix version of equation 6 is:

$$y = \rho W y + \varepsilon \quad (7)$$

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

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

Equation 8 represents the augmented model to investigate the effect of the network based on Equation 7:

$$y = \alpha l_n + \rho W y + gEC + \gamma YSM + \beta X + \varepsilon \quad (8)$$

Where y is the vector of individual immigrant hourly earnings, X is a vector of socio-economic variables, Wy is the ethnic-spatial lag vector of the dependent variable that indicates the first-order ethnic-spatial relationship among individuals, and the coefficient ρ measures the ethnic–spatial autocorrelation of the logarithm of hourly earnings. Thus, ρ indicates the size of the effect of the network in a specific region on earnings. l_n is a vector of ones associated with the parameters α and β. EC represents ethnic concentration, and the parameter vector γ will reveal its effect. YSM is a vector of years since migration; X represents other explanatory variables controlling for conventional human capital and personal characteristics, and year fixed effects. Equation 8 presents the general form of the model. Further details of variables and their specifications are discussed in the section on the data and variables (Sections 4.2 and 4.3).

Verification of SAR model specification

Spatial dependence test:

After identifying the ethnic networks for individual immigrants, we have conducted a series of specification tests for verifying the network effect and choosing the best spatial model for modelling the effect of ethnic capital for immigrants in New Zealand case.

The first test we have done is Moran's I test for spatial autocorrelation for immigrants' earnings. The test result (Moran's I (error)=5.282 (p-value=0.000), MI/DF=0.1042), confirms that the immigrants' earnings are correlated according to their ethnic networks, and the conventional OLS model of immigrants' earning may overlook the ethnic network effect. We observe a highly significant positive ethnic network effect in immigrants' earning, in this case, the conventional model without the spatial spill-over component (e.g. ethnic network in our case) may over-estimate the effect of the included controlled / explanatory variables (LeSage and Pace, 2009).

The second test we have done is the Lagrange Multiplier (LM) test. Moran's I test answers the question whether or not the spatial spill-over effect (or spatial dependence) has been mis-specified by the conventional model; but it does not answer another important question that the spatial autocorrelation is generated by an error process or

omitted spatially lagged dependent variable or both (Anselin and Kelejian, 1997). As a result, we use the LM test for spatial models selection as guided by Anselin (2005). In our case, both the LM-lag and the Robust LM-lag tests reject the null, while the Robust LM-Error failed to do so[6], which suggests that the spatial lag model (SLM) which is also known as spatial autoregressive model (SAR) is preferred (refer to Anselin, 2005).

GMM estimation

We apply the GMM to account for the endogenous spatial lag, based on all exogenous variables *X*, and the corresponding first-order spatial lag of all exogenous variables WX as the set of instruments to explain the spatial lag Wy, as suggested by Anselin (1999). We show that this approach works well in this setting.

LeSage and Kelley (2009 and 2010) discussed the biases of ordinary least squares (OLS) estimation on the parameter ρ in the spatial autoregressive (SAR) model. The maximum likelihood (ML) estimation provides an efficient approach to estimating the parameter of the spatial lag ρ, but it is very computationally complex, especially for a large data set (refer to Anselin 1988). 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 the spatial lag, and for estimating ρ, the parameter of spatial lag. As Kelejian and Robinson (1993) have illustrated, all exogenous variables (X), their spatial lags (WX), and higher spatial lags (e.g., $W^2X,\ W^3X, \ldots, W^nX$) work jointly as a set of instruments for the endogenous spatial lag Wy.

Based on 2SLS estimation, Lee (2007, p. 1) provided evidence that the "GMM estimator (GMME) may be asymptotically more efficient than the two-stage least squares (2SLS) estimator and may be asymptotically as efficient as the ML estimator (MLE)". He indicated that the GMM estimation for those models can be computationally simpler than the ML or quasi maximum likelihood (QML) methods in a general setting (Lee 2007).

[6] The Lagrange Multiplier test results are as follows: Lagrange Multiplier (lag)=28.61 (p-value=0.000); Robust LM (lag) =8.022 (p-value=0.005); Lagrange Multiplier (error)=24.635 (p-value=0.000); and Robust LM (error)= 0.048 (p-value=0.827).

Moreover, and importantly, the GMM estimator corrects for heteroscedasticity.[7] A dynamic spatial lag of order-one (i.e. based on the network of members residing in the same geographic area) is generally applied for socio-economic factors (e.g., Baltagi et al. 2017; Piracha et al. 2016). For example, Conley and Topa (2002) studied spatial patterns in unemployment by analysing agents' order-one social networks. Order-one lag is reasonable for socio-economic factors, where the closest network of resources tends to influence group-member outcomes most prominently.[8] This is in contrast to health sciences, where spread of infection, for example, is typically best represented by spatial lags of higher orders. In this study immigrants from the same country of birth, residing in the same area in the same year, are considered to be order-one group members. We have selected the first-order spatial lag of all exogenous variables WX as the set of instruments, in addition to explanatory variables X, to explain the ethnic spatial lag Wy, as suggested by Anselin (1999).[9]

We further investigate the network effect for three pooled sub-samples for major immigrant country-of-birth groups in New Zealand, and residing within or outside of their group's high concentration areas. The ethnic network may vary across ethnic groups, and it is expected to have a greater economic impact on immigrants residing within high ethnic concentration areas for their ethnic group. We test these hypotheses in this paper.

4 Data

In the following sub-sections we provide a brief discussion of the immigration policy in New Zealand (4.1); followed by a discussion of the data set (4.2); and variable specifications (4.3).

4.1 Background to New Zealand immigration

New Zealand has long been a country of immigrants. The population size, as well as the structure of its labour market, has been significantly affected by immigration. During the past two and a half decades, a series of reforms and adjustments has been made to New

[7] The results of Breusch–Pagan/Cook–Weisberg test for heteroscedasticity could not reject the null hypothesis in our cases (Breusch and Pagan 1980).
[8] Baltagi et al. (2017) similarly examined the network effects on labour contracts of internal migrants in China by using proxies of social networks of order-one.
[9] The full set of instruments for endogenous spatial lag is: $H = [X, WX, W^2X, W^3X, \dots, W^nX]$

Zealand's immigration policy. These reforms increased the probability of prospective immigrants from non-English-speaking regions being accepted. In particular, the Immigration Amendment Act of 1991 and a series of reforms after 1991 reformed New Zealand's immigration system and its selection strategy by selecting applicants based on points for personal characteristics such as education and work experience, regardless of country of birth. In 2003, this strategy was supplanted by the "skilled migrant category" policy, which also incorporated a partner's skills and work experience.

These changes in immigration policy created profound effects on immigrants' labour market performance and the ethnic composition of immigrants (Poot 1998). Using NZIS data, Stillman and Maré (2009) observed that quality changes, measured in education and skills, occurred across different cohorts of immigrants. For example, on average, immigrants who arrived in New Zealand after 1998 have more years of education and are more likely to be employed than those who arrived before 1998.

4.2 Data set

We employ microdata for 10 years (2001–2010) from the NZIS. The NZIS is a rich data set that includes extensive information on employment and earnings. This data set utilises surveys on a randomly selected sample of the labour force each year. The first NZIS was conducted in June 1997 as a supplement to the Household Labour Force Survey for the June quarter (April to June), and it has been conducted annually since 1997. Based on the New Zealand censuses, this data set maintains a random representative selection of the labour force.

A merged microdata set created from the 2001 to 2010 NZIS data sets has been used in this study. The sample we focus on consists of full-time employed male immigrants between 20 and 55 years of age.[10] The pooled panel of cross-sectional data contains 54,422 observations (42,259 New Zealand born and 12,163 immigrant (born overseas) observations). We augment our data with matched spatial data on ethnic concentration from the 1996 and 2001 population censuses.

[10] In this paper, we have focused on the sample of male full-time employed workers to trial our spatial autoregressive model for a relatively more homogenous group of workers, given the complexity of modelling requirements. This research area, however, lends itself to extensions in future studies that also incorporate and address female employment choices.

The dependent variable of interest is the actual hourly wage from the main job in the year of the survey in constant 2006 dollars. This variable serves as an indicator of an individual's labour market performance. In order to adjust for regional income differences, we have converted the individual's hourly wage into the corresponding relative wage, equal to the individual's consumers price index (CPI) adjusted wage rate divided by the regional CPI-adjusted average wage.[11]

The NZIS and the New Zealand census provide classifications linked to ethnicity based on the individual's country of birth. We also use this classification.

The ethnic-spatial lag variable, the explanatory variable of main interest, is derived (as specified in Section 3) across five major country-of-origin groups, by region of residence in New Zealand, and for each year. Based on considerations of geography and language, and ethnic classifications in the NZIS data, immigrants are grouped according to their birthplace into four groups: UK and Ireland; Asia; Pacific Islands; and other countries. We acknowledge that these country-of-birth groups are broad, as constrained by our data set. Nonetheless, the country-of-birth groups UK and Ireland, Asia, and Pacific Islands are numerically and culturally distinct groups, together accounting for 56.3% of the immigrant sample meeting the objectives of our study. This type of classification is in keeping with earlier seminal studies across immigrant ethnic groups (e.g., Borjas 1985, 1995). The data provide variation in this measure for each individual by country-of-birth group, location and year of data. Overall, we believe that the advantages of using the data for a trial of our model outweigh the limitations of the data.

There are 16 distinct locations in our analysis, as determined by the 16 local (regional) government areas. The majority of earlier studies (e.g., Beckhusen, et al. 2013; Borjas 1995; Fong and Jing 2011) examined immigrants' geographical locations in metropolitan statistical areas (MSAs), which are generally organised around a major city. We note that the regional setting in our analysis, which includes one or more major cities with travel duration generally within one to two hours, and is mainly governed by local

[11] We also used the wage deflator as an alternative to the CPI. The results are not sensitive to the choice.

councils, is equivalent to those studies in terms of size.[12] Immigrants living in one suburb of a region (e.g., Auckland) are able to access knowledge/information, indirectly, regarding job opportunities in other suburbs of that region, and a network is likely to develop within that area.

We note that the immigrants from the same country-of-birth group who reside in the same location in the same year may not necessarily know the other members from the same origin in the sample. But as a randomly selected sample, these groups represent the characteristics of the population of the immigrants from the same country of birth who reside in the same locality.

The regional choice of our analysis is guided by the geographic breakdowns of the data set, based on official New Zealand regional specifications. However, the method applied lends itself to applications in future studies to smaller spatial units, such as city or small town boundaries.

We aim to capture the country-of-birth group impact of a regional network for individual members residing in that region. For example, a common scenario is that some immigrants may live in one suburb, yet work in another suburb. Others may have found employment through a friend in a different suburb. Therefore, in considering immigrant earnings, the regional setting of our data helps to best address outcomes in this labour market scenario.

From a statistical point of view, despite the relatively small number of 16 regions, the regions provide sufficient variation for analysis. Specifically, they are significantly different in their economic, demographic, and geographic characteristics. For example, the Auckland region has the highest gross domestic product (GDP) among all New Zealand regions – it accounted for 35.3% of New Zealand's GDP in the year ended March 2013 (Statistics New Zealand 2013). In addition, the skill structure of the labour force and the economic specialisation vary significantly across New Zealand regions. This is a very positive feature of the data from a statistical point of view.

[12] It is common for New Zealand regions to be governed by local councils. The most populated region is Auckland, centred around one metropolitan area, with a population of 1.4 million in the 2013 Census.

Notably variations in variables, including the network and ethnic concentration variables, results from notable differences within regions by country-of-birth group and time.

Table 1 presents socio-economic characteristics for people born in New Zealand and for the three most populous immigrant groups in our analysis. Table A1 in the Appendix provides definitions of the variables.

The data set includes immigrants from diverse countries of birth with significant differences across groups in their economic characteristics. In general, Asian immigrants are younger, are much more likely to hold a university degree, and have resided in New Zealand for fewer years compared to immigrants from the UK and Ireland and the Pacific Islands. We also note that the average hourly wages of immigrants born in Asia and the Pacific Islands are on average lower than for immigrants born in the UK and Ireland.

In our sample, 41.2% of the immigrants arrived in New Zealand between 1992 and 2003. It is worth noting that more than half (62%) of the Asian immigrants arrived in New Zealand during this time period. The "years since migration" of immigrants from the UK and Ireland is greater than 16 years and is the highest compared to other immigrant groups. Immigrants in the data set follow national averages, and there is a wide range of years since migration providing desirable variation for the analyses.

Table 1. Socio-economic indicators among immigrant ethnic groups

		Selected country-of-birth groups		
	Pooled Immigrants	**UK & Ireland**	**Asia**	**Pacific**
Hourly earnings (mean)	23.08	25.65	20.48	19.15
University degree (%)	29	29.22	45.36	17.56
Age (mean)	37.35	39.63	34.81	36.22
Years since migration (mean)	12.05	16.66	8.74	10.81
Observations	12,163	2,464	1,466	2,921

Notes: NZIS (2001–2010), employed males aged 20–55; hourly earnings are in constant 2006 dollars

4.3 Variable specifications

As discussed in detail in Section 3.1, the spatial ethnic network variable Wy is the product of the W matrix applied on the vector y of the ethnic group spatial earnings derived for each year of data. Matrix W and vector y are derived from entries for 36 country-of-birth groups of immigrants, by 16 regions, for a year of data. Thus, the weight matrix incorporates different spatial weights for different regions and country-of birth groups of migrants and year.

Vector y on earnings is of special interest, as it represents the economic outcome of labour market engagement. earnings are expected to also serve as a proxy for other forms of group income. Internationally, earnings are the largest component of household income and national income generated in each country. For example, in New Zealand two thirds of all household income is generated in the form of wages and salaries (NZ Parliament, 2018). In addition, earnings of groups have a high correlation with other income and economic resources of the group. Therefore, in addition to representing purchasing power, higher group earnings reflect networks of resources for the group. Examples are availability of loanable funds, and referrals and trusted information for job seekers and employers in the network for jobs with higher earnings.

For the measurement on a respondent's ethnic-spatial concentration (EC_{cr}), we merge the NZIS data with New Zealand census data for the years 1996 and 2001. The ethnic concentration variable is a lagged measure from the census before the last one for the year used in the study. This variable utilises the official New Zealand census tables, and it is derived according to the conventional formula $EC_{cr} = \frac{Pop_{\,cr\,(t-k)}}{Pop_{r(t-k)}}$, where "$c$" denotes country-of-birth group (36 groups), "r" represents a specific region (16 regions), "t" represents the year of data (10 years), and "k" represents the time lagged nature of the variable based on the previous census.

It is relevant to control for ethnic concentration in this setting, although interpreting the coefficient of this variable as a causal effect requires some strong

assumptions. For example, there could be positive or negative sorting. Considering this, we are comfortable with interpreting this coefficient as an association between ethnic concentration and an immigrant's earnings. We further take the following measures. First, we follow Edin et al. (2003) by incorporating a lagged value of ethnic concentration based on an earlier five-yearly census. For example, for the years 2001 to 2005 of our data, we use ethnic concentration based on the 1996 census. Similarly, we use ethnic concentration based on the 2001 census for our data in the years 2006 to 2010. Second, we report results with and without this variable (Section Five, Table 2), which confirm that our coefficients of interest (i.e. for network effects, returns to educational qualifications, and years since migration), do not change significantly with the inclusion of this variable. Third, we report results for sub-samples that reside within or outside of ethnic enclaves (Section Five, Table 3), hence removing the need to account for ethnic concentration within the model. In the next section we present our results. In the econometric models the natural logarithm of EC_{cr} is incorporated.[13]

The other explanatory variables measure conventional human capital factors (educational qualifications, and years of experience); personal characteristics (i.e. marital status, and broad ethnic-groups binary variables); cohort of immigrant arrivals; and year fixed effects. Table A1 (Appendix) provides the definition of variables. As we show in the results section, these variables have the expected signs and perform statistically well.

In our analysis, we also incorporate an alternative dichotomised measure of ethnic concentration, which signifies residing in major enclave region (high ethnic concentration (HEC) areas), based on conventional residential concentration quotient (RCQ) measures. We discuss the added variable specifications in detail in the discussion of extended analyses by enclave region (Section 5.4).

[13] Goodness of fit tests confirm the logarithmic specification of this variable in the econometric model.

5 Results

In this section, we first show that our data are comparable to studies from other Western immigrant-receiving countries in a conventional model of immigrant earnings. Then we examine whether ethnic-spatial economic networks have a significant effect on immigrant earnings, and whether the effect is positive or negative.

5.1 Conventional model of immigrant earnings

The results of the conventional model – reported in Table 2 (column 1) and based on an OLS estimation of the logarithm of hourly earnings – confirm a number of general findings. The hourly earnings of immigrants are affected positively by educational degrees and years of work experience. Among immigrants, each year since migration has an additional positive and significant impact on earnings.

In addition, there are significant differences in earnings by ethnic group, even after controlling for conventional human capital factors. Immigrants from both Asia and the Pacific Islands show a remaining hourly earnings difference of over 15% compared to immigrants from the UK and Ireland. (Supplementary conventional models that compare immigrant earnings to the New Zealand-born population show a remaining 20% disadvantage for immigrants born in Asia and the Pacific Islands. Immigrants from the UK and Ireland, in comparison, have a much smaller earnings disadvantage of under 4%). These results are available in Table A3 in the Appendix.

There is also strong evidence of cohort effects among immigrants. As expected, immigrants who arrived after 1992 perform significantly better than the earlier cohorts, who arrived prior to the skilled migration policy changes of 1991. In addition, the most recent group of immigrants, who arrived after the fine-tuning of the skilled migration policy in 2003, perform significantly better than all previous groups.

We augment our conventional immigrant earnings model by including the two spatial measures of ethnic networks (of economic resources) and ethnic concentration.

These results are summarised in Table 2 (columns 4 and 5) and are discussed in the next section.

5.2 Spatial model with ethnic network effect

Immigrant ethnic networks are expected to be affected by the strength of the resources, as reflected by the group members' average earnings, for the ethnic group in their spatial setting.

Table 2 (spatial models with network variables) shows the effects of two country-of-origin group network variables. Column 4 shows the impact of the ethnic-spatial network variable on immigrants' real hourly wage rate. Column 5 shows the added effect of ethnic concentration. As discussed in Section 3, OLS provides an inconsistent estimation on the coefficients for the SAR (also known as SLM) model adopted in this analysis. Therefore, these models are estimated by GMM (columns 4 and 5) to address endogeneity of the network effect variables. We also find that the OLS estimation (columns 2 and 3) tends to underestimate the coefficient of ethnic spatial economic network variable (spatial lagged dependent variable) in our case. Importantly, both of the spatial models indicate a strong and significant general ethnic network effect on immigrants' labour market performance. The coefficients for both ethnic network and ethnic concentration effects are positive and significant. These results are robust across variations of the model and despite controlling for relevant human capital factors.

Additionally, in comparison to the conventional model, the spatial model performs significantly better statistically.

The results in Table 2 establish a positive and strong correlation between the strength of economic resources of the regional ethnic networks and earnings outcomes for these immigrants. As Table 2 reports (columns 4 and 5), a 10% increase of the average hourly wage of (the other members of) the ethnic-spatial group at the mean is associated with an approximately 0.6% spill-over increase in a member's hourly wage.[14]

[14] For example, for a coefficient of 0.065, a 10% increase in the network variable represents a change in Y = [(1.10) ^ 0.065] – 1 = 0.0062; and for a coefficient of 0.062, a 10% increase in the network variable represents a change in Y = [(1.10) ^ 0.062] – 1 = 0.0059, approximating 0.6%.

These results are in contrast to Edin et al. (2003), who found an earnings loss, at sample means, of 1.9% from a standard deviation increase in ethnic concentration, and positive effects for the higher end of the distribution (Edin et al. 2003, p. 321). We find positive and significant effects at the mean. One difference, in addition to the different estimation methods, between their study and ours is that they focused on a sample of refugees in Sweden, and the current study is based on all immigrants to New Zealand. As such, our results are consistent with positive economic network effects.

In addition, living in areas where a larger ethnic concentration has been present in the recent past is also associated with additional positive earnings outcomes for immigrants across these ethnic groups. As shown in Table 2 (column 5), a 10% increase in ethnic concentration ratio is associated with a 0.23% increase in earnings (coefficient of 0.024). This result is also consistent with positive opportunities created for these groups of immigrants from the same country of birth who reside in the same locations.

Table 2. Immigrant earnings with network effect: Natural log of hourly wage

	Conventional model	**Spatial Model with Network Variable**			
	OLS Estimation	OLS Estimation		GMM Estimation	
	(1)	(2)	(3)	(4)	(5)
Ethnic-spatial economic network	/	**0.015***** (0.006)	**0.023***** (0.008)	**0.065***** (0.010)	**0.062***** (0.010)
Ln Ethnic concentration (lagged)	/	/	**0.014***** (0.001)	/	**0.024***** (0.008)
		/		/	
Human Capital					
Experience	**0.037***** (0.002)	**0.036***** (0.002)	**0.036***** (0.002)	**0.036***** (0.002)	**0.036***** (0.002)
Experience -squared	**−0.001***** (0.00004)	**-0.001***** (0.0001)	**-0.008***** (0.0001)	**-0.001***** (0.0001)	**-0.008***** (0.0001)
High School Qualification	**0.174***** (0.025)	**0.180***** (0.023)	**0.180***** (0.023)	**0.181***** (0.028)	**0.181***** (0.028)
Vocational Qualification	**0.191***** (0.011)	**0.184***** (0.011)	**0.186***** (0.011)	**0.193***** (0.011)	**0.195***** (0.011)
University Degree	**0.423***** (0.011)	**0.412***** (0.011)	**0.413***** (0.011)	**0.428***** (0.013)	**0.430***** (0.013)
Other Post School Qualification	**0.162***** (0.017)	**0.158***** (0.017)	**0.160***** (0.016)	**0.163***** (0.021)	**0.166***** (0.021)
Personal Characteristics					
Years Since Migration (YSM)	**0.012***** (0.002)	**0.011***** (0.002)	**0.011***** (0.002)	**0.011***** (0.002)	**0.010***** (0.002)
YSM-squared	**−0.0002***** (0.00003)	**-0.0002***** (0.00004)	**-0.0002***** (0.00004)	**-0.0002***** (0.00004)	**-0.0002***** (0.00004)
Married	**0.079***** (0.01)	**0.073***** (0.011)	**0.074***** (0.010)	**0.070***** (0.012)	**0.071***** (0.012)
UK and Ireland	**0.087***** (0.012)	**0.076***** (0.012)	**0.070***** (0.012)	**0.080***** (0.015)	**0.067***** (0.015)
Asia	**−0.09***** (0.013)	**-0.086***** (0.013)	**-0.095***** (0.013)	**-0.084***** (0.016)	**-0.091***** (0.016)
Pacific Islands	**−0.084***** (0.011)	**-0.072***** (0.011)	**-0.072***** (0.011)	**-0.083***** (0.013)	**-0.085***** (0.013)
Other countries	Ref	Ref	Ref	Ref	Ref
Cohort effect					
Arrived 1992-2003	**0.053***** (0.019)	**0.043**** (0.022)	**0.043**** (0.019)	**0.049***** (0.022)	**0.048***** (0.022)
Arrived 2003-2010	**0.111***** (0.028)	**0.100***** (0.028)	**0.100***** (0.028)	**0.100***** (0.033)	**0.098***** (0.033)
Year Fixed Effect	Yes	Yes	Yes	Yes	Yes
Observations	12163	12163	12163	12163	12163
F statistics	164.54	185.48	177.06	163.90	158.01
Adjusted R-square	0.244	0.246	0.240	0.251	0.251

Notes: (1) Employed male immigrants, aged 20–55, NZIS (2001–2010). (2) Standard errors in parentheses. (3) *$p < 0.10$, **$p < 0.05$, ***$p < 0.01$. (4) Spatial models with network effects (results in columns 4 and 5) are estimated by GMM to address endogeneity of the network.

5.3 Test of exogeneity of key explanatory (X) variables.

As discussed in Section 3, we conducted additional auxiliary tests to verify the validity of our modelling choice and the interpretation of the results. These tests included tests of exogeneity of the explanatory variables.

Since some of the independent variables in our model might be potentially endogenous such as educational attainment variables, years since migration (YSM) and (potential) work experience, we conducted the Wu-Hausman F and Durbin-Wu-Hausman Chi-square tests for these variables. All three test results have p-values such that the hypothesis of exogeneity cannot be rejected at the 5% and 1% significance levels. For YSM, the p-values for both tests are around 0.21. For Potential work experience both p-values are around 0.38, and for educational attainment variables the p-values are at around 10%. These analyses confirm the validity of our approach and our main results that are reported in Table 2.[15]

5.4 Results by enclave residence and country-of-origin group

In this section, we examine this question further by taking a closer look at the earnings of immigrants who live within, or outside of, regions that contain enclaves for their ethnic group. We examine the effect of ethnic economic networks in the spatial setting for each of the three sub-samples of immigrant ethnic groups of interest in our analysis. This additional analysis allows us to examine if our earlier results are robust across locality and ethnic groups.

The effect of the ethnic network may vary due to the culture, norms, and other characteristics of different ethnic (country-of-origin) groups. We expect that immigrants

[15] Kelejian and Piras (2014) propose a test to examine the endogeneity of the weight matrix (W) itself. Following their approach, we test and show elsewhere (*Papers in Regional Science*) that this potential endogeneity is not a concern in this analysis.

who live in high concentration areas of their ethnic group experience a greater impact from their ethnic network compared to immigrants who live outside of their ethnic group's high concentration areas. For example, previous research has shown that immigrants may find greater opportunities of employment through geographic concentration, as ethnic networks may assist new immigrants to adjust to the new environment in the destination country (e.g., Edin et al. 2003; Munshi 2003; Piore 1979).

In this section we extend our analysis to further examine the impact of ethnic resources on the hourly earnings of immigrants.[16] This added specification provides an opportunity to compare the impact of the ethnic network for immigrants who reside within, compared to those who reside outside of, relatively high concentration areas for their immigrant ethnic group. We report results for three pooled major sub-samples of immigrant ethnic groups: UK and Ireland; Asia; and Pacific Islands.

The results for immigrants living within, and outside of, their ethnic group's major enclave areas are presented in this section. For this added analysis, we define major enclaves as high ethnic concentration (HEC) areas, based on the conventional calculations of a residential concentration quotient (RCQ), as defined in Equation 9 below. An RCQ is conventionally used in labour market research to dichotomise the continuous ethnic concentration (EC_{cr}) measure according to whether or not the concentration of population from the same ethnic background is higher than a threshold level. Equation 9 provides a ratio:

$$RCQ_c = \frac{Pop_{cr}}{Pop_r} \Big/ \frac{Pop_{cm}}{Pop_m} \tag{9}$$

where r = (1,…,n) represents different localities; Pop_{cr} is the population from a typical immigrant group (c) residing in location r; Pop_r refers to the total population in location r; Pop_{cm} is the total population size of current and earlier immigrants from immigrant group c in the host country; and Pop_m is equal to the host t country's total population size. Therefore, when the RCQ is equal to 1, ethnic concentration of immigrant ethnic group 'c' in a certain region 'r' is on par with the country average, and an RCQ larger than 1

[16] More details about categorising New Zealand regions based on ethnic concentration are provided in Table A2 in the Appendix.

shows a greater level of ethnic concentration for immigrant group *c*. In this paper, we adopt a threshold of 1.2 for the RCQ to identify the primary ethnic enclaves for immigrants in the country. The threshold value of 1.2 allows us to identify at least one HEC area for every ethnic group in our analysis.[17] Table A2 in the Appendix reports our derived RCQ measures for the three major immigrant groups of interest in our analysis, for all New Zealand regions. Different studies adopt different threshold values appropriate to the data for defining an HEC area. For example, Zhu et al. (2013) adopt a value of 1.5 for the RCQ in their study.[18]

According to this definition and the derived RCQ regional measures, the Auckland region is identified as an HEC area for Asian immigrants, and it is the only locality that meets the threshold requirement for this immigrant group. Wellington and Nelson regions are HEC areas for immigrants from the UK and Ireland. This shows it is possible for more than one primary ethnic enclave to exist for an immigrant group.

Table 3 summarises results for the six sub-samples of immigrants who live within or outside of their ethnic group's primary enclave (HEC) area(s). To highlight the findings, we consider the results for the three major immigrant ethnic groups: UK and Ireland; Asia; and Pacific Islands. Based on our ethnic resources network hypothesis, we expect that the ethnic network variable would exert a greater impact for immigrants who reside within their ethnic groups' high ethnic concentration area(s).

The results in Table 3 confirm that for all three groups, the ethnic network variable has a larger and more significant coefficient for immigrants residing within, compared to the same group residing outside of, their major enclave areas. These disaggregated results further confirm that, for all three major immigrant ethnic groups, the ethnic network variable has a positive and highly significant effect in the enclave areas. These results further support our earlier results on the co-dependence of immigrant ethnic group earnings and ethnic concentration, and the greater impact of ethnic networks in areas where ethnic concentration is greater.

[17] In Parks' (2004) paper the cut-off RCQ was 5, for example, based on their data.
[18] We conducted auxiliary tests based on size of the enclave, and RCQ with threshold levels of 1.1 and 1.3. Our main results are not sensitive to this selection.

For the Asian group of immigrants, the highly significant effect within the enclave (HEC) area becomes insignificant outside of the enclave. For those who live in the enclave, a 10% increase of the average hourly wage of the ethnic-spatial group is associated with an approximately 4% (3.9%) spill-over increase in a member's hourly wage. For immigrants born in the Pacific Islands, the coefficient is about tenfold within the HEC area, compared to outside of the HEC area. For this group living within the enclave, a 10% increase of the average hourly wage of the ethnic-spatial group is associated with an approximately 1% (0.9%) spill-over increase in a member's hourly wage. These results strongly support our expectation that the ethnic network effect is important and significant.

Table 3. Natural log of hourly wage with network effect
by enclave residence and country-of-origin group (GMM estimation)

	Residing within enclave regions			Residing outside of enclave regions		
	UK & Ireland	Asia	Pacific	UK & Ireland	Asia	Pacific
Ethnic-spatial economic network	0.119**	0.093*	0.401***	0.087***	0.003	0.045***
Observations	511	855	1901	1953	611	1020

Notes: (1) Residing within enclave refers to immigrant residence in a high ethnic concentration region (see Section 5.4). (2) The models include other variables as in Table 2. The full set of results is available in Table A4 in the Appendix. (3) *$p < 0.10$, **$p < 0.05$, ***$p < 0.01$. (4) Employed male immigrants, aged 20–55, NZIS (2001–2010).

For immigrants from the English-speaking origins of the UK and Ireland, the impact is also positive and significant (an approximately 1.3% increase for a 10% increase in the group's average earnings). However, the difference of the impact from living within or outside of the enclave is relatively smaller.

6 Conclusion

The thrust of this paper is that while immigrants may face many barriers to employment in the host country's labour market, ethnic network and ethnic concentration are positively associated with immigrant earnings.

We have enhanced the conventional migrant economic performance model by incorporating a spatial autoregressive network approach of immigrant ethnic capital across locality and time. An important feature of this approach is that it relaxes the independent and identically distributed (i.i.d.) assumption of immigrants' earnings, incorporating the potential co-dependence of immigrant ethnic groups' economic outcomes. It also incorporates the quality (strength) of ethnic group resources. Previous international studies generally adopt ethnic concentration, or the size of the group, as a proxy of immigrants' network in the host country. We have explored an alternative approach, and we find conclusive results.

Our results are consistent with positive network effect hypotheses. Incorporating two different measures on ethnic group economic resources and ethnic-spatial concentration with GMM estimation, we find positive and significant spill-over impacts from both ethnic-spatial networks and ethnic concentration on immigrants' earnings. We find that their ethnic group spatial concentration and resources are positively and significantly associated with immigrants' wage growth. A 10% increase of the average hourly wage of the ethnic-spatial group at the mean is associated with an approximately 0.6% spill-over increase in a member's hourly wage.

The results further show that the network variable associations with the immigrants' earnings are significantly greater when immigrants reside in areas of high ethnic concentration of their country-of-birth group. A 10% increase of the average hourly wage of the ethnic-spatial group is associated with an approximately 4% spill-over increase in a member's hourly wage if living within the enclave for Asian immigrants; approximately 1.1% for immigrants from the UK and Ireland; and approximately 1% for immigrants from the Pacific Islands. An added result is that the marginal gains from the network when residing within the enclave are proportionately greater for immigrants from non-English-speaking country-of-birth groups. These results support the hypothesis that the positive economic network effects supersede potential negative integration effects in this analysis. This effect is notably stronger for Asian immigrants, who have higher proportions with university degrees, compared to the immigrants from the Pacific Islands.

Finally, we find that the spatial autoregressive network approach adopted in this paper provides opportunities for labour economics research that enhances the conventional immigrant economic performance and assimilation/integration analyses.

Appendix

Table A1. Definitions of variables

Hourly Wage	Dependent variable: Natural log transferred hourly wage.
Ethnic-Spatial Economic Network (W ln(y))	The weighted ethnic-spatial lag (The weighted (average) natural logarithm of hourly wage (W ln(y)) of an individual's country-of-birth spatial network (excludes the respondent's wage), by location (16 regions), country of birth (36 country-of-birth groups), and year (10 years of data)). W ln(y) represents the ethnic–spatial autocorrelation of the natural logarithm of hourly earnings.
Ethnic Concentration (EC_{cr})	The natural log of the lagged proportion of the population of a specific ethnic group (36 country-of-birth groups) to the total population size in a specific region (16 regions) from the census previous to the last census by year of data (Refer to Section 4.3).
Years Since Migration (YSM)	Years of residence in the host country.
Married	Binary variable=1 if married.
Education Attainment	Four binary variables=1 if that individual's highest education attainment is either high school certificate, vocational certificate, other post school certificate, or university degree; base group is under high school.
Country-of-Origin Group	Three binary variables = 1 if that individual's country of birth is either Asia; UK and Ireland; or Pacific Islands. The base group is rest of world.
Cohort effects	Two binary variables = 1 if arrived between the years 1992–2003 (after the first migration policy change in 1991) or 2004–2010 (after policy fine-tuning of 2003); base group arrived before 1992.
Year fixed effects	Nine binary variables = 1 if the observation comes from survey years 2002, 2003, 2004, … 2010 (year 2001 is the base group).
Residential Concentration Quotient (RCQ)	RCQ is used to derive a dichotomous measure of the continuous ethnic concentration (ECcr) measure according to whether or not the concentration of population from the same country of birth background is higher than a threshold level. (Refer to Section 5.4).

Table A2. Residential concentration quotient (RCQ) and high ethnic concentration (HEC) areas

Country of origin	Regions							
	1	**2**	**3**	**4**	**5**	**6**	**7**	**8**
UK and Ireland	0.960	1.086	0.865	1.002	0.523	0.878	0.765	0.770
Asia	0.185	2.044*	0.524	0.350	0.149	0.246	0.233	0.401
Pacific Islands	0.198	2.271*	0.376	0.237	0.264	0.415	0.133	0.299
Country of origin	**Regions**							
	9	**10**	**11**	**12**	**13**	**14**	**15**	**16**
UK and Ireland	1.208*	1.195	1.408*	0.962	0.588	1.030	0.921	0.492
Asia	0.818	0.158	0.306	0.188	0.121	0.733	0.491	0.131
Pacific Islands	1.008	0.068	0.197	0.171	0.105	0.291	0.205	0.150

Notes: (1) * High ethnic concentration (HEC) areas for each ethnic (country of birth) group are based on RCQ above the threshold of 1.2.

(2) RCQ measures are derived based on Equation 9 in the paper and raw population data from the published 2006 New Zealand Census table: Birthplace (Broad Geographic Areas) for the Census (Usually Resident Population Count (Statistics New Zealand, 2006)).

(3) New Zealand regions are: [1] Northland Region, [2] Auckland Region, [3] Waikato Region, [4] Bay of Plenty Region, [5] Gisborne Region, [6] Hawke's Bay Region, [7] Taranaki Region, [8] Manawatu-Wanganui Region, [9] Wellington Region, [10] Tasman Region, [11] Nelson Region, [12] Marlborough Region, [13] West Coast Region, [14] Canterbury Region, [15] Otago Region, and [16] Southland Region.

(4) Sample: employed male immigrants, aged 20–55, NZIS (2001–2010).

Table A3. Conventional model of immigrant natural log of hourly wage

	Immigrants	NZ-Born	Pooled Sample
Human Capital			
Experience	0.037***	0.033***	0.033***
	(0.002)	(0.001)	(0.001)
Experience-squared	−0.001***	−0.001***	−0.001***
	(0.00004)	(0.00002)	(0.00002)
High School Qualification	0.174***	0.181***	0.179***
	(0.025)	(0.01)	(0.009)
Vocational Qualification	0.191***	0.193***	0.193***
	(0.011)	(0.005)	(0.004)
University Degree	0.423***	0.524***	0.492***
	(0.011)	(0.006)	(0.005)
Other Post School Qualification	0.162***	0.174***	0.174***
	(0.017)	(0.011)	(0.009)
Personal Characteristics			
Years Since Migration (YSM)	0.012***	/	0.006***
	(0.002)	/	(0.001)
YSM-squared	−0.0002***	/	−0.0001***
	(0.00003)	/	(0.00003)
Married	0.079***	0.125***	0.116***
	(0.01)	(0.005)	(0.004)
UK and Ireland	0.087***	/	−0.038*
	(0.012)	/	(0.02)
Asia	−0.09***	/	−0.226***
	(0.013)	/	(0.021)
Pacific Islands	−0.084***	/	−0.208***
	(0.011)	/	(0.02)
Rest of World	/	/	−0.132***
	/	/	(0.02)
Cohort Effect			
Arrived 1992–2003	0.053***	/	−0.007
	(0.019)	/	(0.014)
Arrived 2003–2010	0.111***	/	0.017
	(0.028)	/	(0.019)
Year fixed effects	Yes	Yes	Yes
Observations	12163	42259	54422
F Statistics	164.54	764.18	640.73
Adjusted R^2	0.244	0.224	0.227

Notes: NZ-Born refers to New Zealand born.

(1) Standard errors in parentheses. (2) *$p < 0.10$, **$p < 0.05$, ***$p < 0.01$. (3) Employed male immigrants, aged 20–55, NZIS (2001–2010).

Table A4. Natural log of hourly wage with network effect by country-of-origin group, residence within or outside of high group ethnic concentration area (HEC) (GMM estimation)
(*Full set of results corresponding to Table 3*)

	Resident *within* high ethnic concentration (HEC) areas			**Resident *outside of* high ethnic concentration (HEC) areas**		
	UK & Ireland	**Asia**	**Pacific**	**UK & Ireland**	**Asia**	**Pacific**
Ethnic-spatial economic network	0.119**	0.093*	0.401***	0.087***	0.003	0.045***
	(0.051)	(0.055)	(0.097)	(0.020)	(0.033)	(0.015)
Human capital						
Experience	0.039***	0.032***	0.020***	0.035***	0.036***	0.036***
	(0.008)	(0.006)	(0.004)	(0.005)	(0.006)	(0.006)
Experience-squared	−0.0009***	−0.0007***	−0.0004***	−0.0007***	−0.0009***	−0.0008***
	(0.0002)	(0.0002)	(0.0001)	(0.0001)	(0.0002)	(0.0001)
High school degree	0.300**	0.162**	0.143***	0.274***	0.143**	0.063
	(0.123)	(0.070)	(0.042)	(0.099)	(0.058)	(0.074)
Vocational degree	0.134***	0.147***	0.256***	0.144***	0.149***	0.151***
	(0.051)	(0.043)	(0.025)	(0.026)	(0.048)	(0.035)
University degree	0.422***	0.406***	0.370***	0.407***	0.436***	0.465***
	(0.070)	(0.037)	(0.033)	(0.033)	(0.051)	(0.052)
Other post school degree	0.227**	0.082	0.160***	0.098**	0.309***	0.268***
	(0.090)	(0.058)	(0.046)	(0.046)	(0.103)	(0.078)
Personal characteristics						
Years since migration (YSM)	0.024**	0.036***	0.006	0.010*	0.026***	0.022***
	(0.011)	(0.007)	(0.005)	(0.005)	(0.008)	(0.008)
YSM-squared	−0.0004**	−0.0007***	0.00002	−0.0002**	−0.0003**	−0.0003**
	(0.0002)	(0.0001)	(0.0001)	(0.0001)	(0.0001)	(0.0001)
Married	0.100*	0.049	0.020	0.125***	0.095**	0.065
	(0.052)	(0.037)	(0.027)	(0.028)	(0.046)	(0.048)
Cohort effects	Yes	Yes	Yes	Yes	Yes	Yes
Year fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
Observations	511	855	1,901	1,953	611	1,020
F statistics	7.032	17.282	35.298	27.565	12.403	19.910
Adjusted R^2	0.191	0.276	0.265	0.214	0.272	0.271

Notes: (1) Standard errors in parentheses. (2) *$p < 0.10$, **$p < 0.05$, ***$p < 0.01$. (3) Employed male immigrants, aged 20–55, NZIS (2001–2010).

References

Adjemian MK, Lin C-YC, Williams J (2010) Estimating spatial interdependence in automobile type choice with survey data. *Transportation Research Part A: Policy and Practice* 44: 661–675

Anselin L (1988) *Spatial econometrics: Methods and models*. Kluwer Academic Publishers, Dorddrecht

Anselin L (1990) Some robust approaches to testing and estimation in spatial econometrics. *Regional Science and Urban Economics* 20: 141–163

Anselin L (1990) Spatial dependence and spatial structural instability in applied regression analysis. *Journal of Regional Science*. 30:185-207. doi:10.1111/j.1467-9787.1990.tb00092.x

Anselin L (1999) *Spatial Econometrics*. Bruton Center, School of Social Sciences, University of Texas, Dallas

Anselin L (2005) *Exploring Spatial Data with GeoDa: A Workbook*, Center for Spatially Integrated Social Science. http://www.csiss.org/clearinghouse/GeoDa/geodaworkbook.pdf.

Anselin L, Kelejian H H (1997) Testing for spatial error autocorrelation in the presence of endogenous regressors. *International Regional Science Review* 20(1 and 2):153-182

Baltagi BH, Deng Y, Ma X (2017) Network effects on labor contracts of internal migrants in China: A spatial autoregressive model'. *Empirical Economics* https://doi.org/10.1007/s00181-017-1333-3

Battu H, Seaman P, Zeno Y (2011) Job contact networks and the ethnic minorities. *Labour Economics* 18: 48–56

Beckhusen J, Florax RJGM, de Graaff T, Poot HJ, Waldorf BS (2013) Living and working in ethnic enclaves: English language proficiency of immigrants in US metropolitan areas. *Papers in Regional Science* 92(2): 305–328

Beenstock M, Chiswick BR, Paltiel A (2010) Testing the immigrant assimilation hypothesis with longitudinal data. *Review of Economics of the Household* 8: 7–27

Bertrand M, Luttmer EFP, Mullainathan S (2000) Network effects and welfare cultures. *The Quarterly Journal of Economics* 115: 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

Borjas GJ (1985) Assimilation, changes in cohort quality, and the earnings of immigrants. *Journal of Labor Economics* 3: 463–489

Borjas GJ (1992) Ethnic capital and intergenerational mobility. *The Quarterly Journal of Economics* 107: 123–150

Borjas GJ (1995) Ethnicity, neighborhoods, and human-capital externalities. *American Economic Review,* LXXXV: 365–390

Borjas GJ (1999) *Heaven's Door—Immigration Policy and the American Economy,* Princeton University Press, Princeton

Breusch T, Pagan A (1980) The Lagrange multiplier test and its applications to model specification in econometrics. *Review of Economic Studies* 47: 239–253

Chiswick BR (1978) The effect of Americanization on the earnings of foreign-born men. *The Journal of Political Economy* 86: 897–921

Chiswick BR, Lee YL, Miller PW (2005) Immigrant earnings: A longitudinal analysis. *Review of Income and Wealth* 51: 485–503

Chiswick BR, Miller PW (1996) Ethnic networks and language proficiency among immigrants. *Journal of Population Economics* 9: 19–35

Chiswick BR, Miller PW (2002a) *Do enclaves matter in immigrant adjustment?* IZA Discussion Paper 449, IZA

Chiswick BR, Miller PW (2002b) Immigrant earnings: Language skills, linguistic concentrations and the business cycle. *Journal of Population Economics* 15: 31–57

Conley TG, Topa G (2002) Socio-economic distance and spatial patterns in unemployment. *Journal of Applied Econometrics* 17: 303–327

Cutler DM, Glaeser EL (1997) Are ghettos good or bad. *Quarterly Journal of Economics* 112: 827–872

De Jong GF, Madamba, AB (2001) A double disadvantage? Minority group, immigrant status, and unemployment in the United States. *Social Science Quarterly* 82(1): 117–130

Dustmann C, Fabbri F (2003) Language proficiency and labour market performance of immigrants in the UK. *The Economic Journal* 113(489), 695–717

Edin PA, Fredriksson P, Aslund O (2003) Ethnic enclaves and the economic success of immigrants: Evidence from a natural experiment. *Quarterly Journal of Economics* 118: 329–357

Fong E, Jing S (2011) Explaining ethnic enclave, ethnic entrepreneurial and employment niches: A case study of Chinese in Canadian immigrant gateway cities. *Urban Studies* 48(8): 1605–1634

Frijters P, Shields MA, Price SW (2005) Job search methods and their success: A comparison of immigrants and natives in the UK. *The Economic Journal* 115: F359-F376

Fotheringham, AS, Brundson, C, Charlton, M (2002) *Geographically weighted regression: The analysis of spatially varying relationships*. Chichester, West Sussex, England. Wiley.

Goetzke F (2008) Network effects in public transport use: Evidence from a spacially autoregressive mode choice model for New York. *Urban Studies* 45(2): 407–417

Goetzke F, Weinberger R (2012) Separating contextual from endogenous effects in automobile ownership models. *Environment and Planning A* 44: 1032–1046

Goetzke, F, Rave, T (2015) Automobile access, peer effects and happiness. Transportation. 42(5): 791–805

Kalenkoski CM, Lacombe DJ (2013). Minimum wages and teen employment: A spatial panel approach. *Papers in Regional Scienc*e 92(2): 407–418

Kelejian H, Prucha I (1998) A generalized spatial two stage least squares procedure for estimating a spatial autoregressive model with autoregressive disturbances. *Journal of Real Estate Finance and Economics* 17: 99–121

Kelejian HH, Piras G (2014) Estimation of spatial models with endogenous weighting matrices, and an application to a demand model for cigarettes. *Regional Science and Urban Economics 46*:140-149

Kelejian HH, Robinson DP (1993) A suggested method of estimation for spatial interdependent models with autocorrelated errors, and an application to a county expenditure model. *Papers in Regional Science* 72: 297–312

Kossoudji S (1988) English language ability and the labor market opportunities of Hispanic and East Asian immigrant men. *Journal of Labor Economics* 6: 205–228

Lee L (2007) GMM and 2SLS estimation of mixed regressive, spatial autoregressive models. *Journal of Econometrics* 137: 489–514

LeSage J, Kelley P (2010) Spatial econometric models. In: Fischer MM, Getis A (eds), *Handbook of applied spatial analysis: Software tools, methods and applications*. Springer, Berlin

LeSage J, Pace RK (2009) *Introduction to spatial econometrics*. Chapman & Hall/CRC, Boca Raton FL

Logan JR, Alba RD, Stults BJ (2003) Enclaves and entrepreneurs: Assessing the payoff for immigrants and minorities. *International Migration Review* 37: 344–388

Masgoret A, McLeod K, Tausi M, Ferguson B, Plumridge E, Duke T (2012) *Labour market integration of recent migrants in New Zealand.* Department of Labour, Wellington.

McDonald JT, Worswick C (1999) The earnings of immigrant men in Australia: Assimilation, cohort effects, and macroeconomic conditions. *Economic Record* 75: 49–62

Munshi, K (2003) Networks in the modern economy: Mexican migrants in the U. S. labor market. *The Quarterly Journal of Economics* 118 (2): 549–599

New Zealand Parliament (2018) Distributional Analysis: Background Paper Tax Working Group, Wellington

Parks, V (2004) The gendered connection between ethnic residential and labor-market segregation in Los Angeles. *Urban Geography* 25 (7): 589-630.

Piore, MJ (1979) *Birds of passage: Migrant labor and industrial societies*. Cambridge University Press: Cambridge

Piracha, M, Tani, M, Vaira-Lucero, M (2016) Social capital and immigrants' labour market performance. *Papers in Regional Science* 95: 107–126.

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

Portes A (1987) The social origins of the Cuban enclave economy of Miami. *Sociological Perspectives* 30(4): 340–372

Portes A, Shafer S (2007) Revisiting the enclave hypothesis: Miami twenty-five years later. *Research in Sociology of Organizations* 25: 157–190

Statistics New Zealand (2001) *2001 Census Data*. Statistics New Zealand, Wellington

Statistics New Zealand (2006) *2006 Census Data*. Statistics New Zealand, Wellington

Statistics New Zealand (2013) *Regional Gross Domestic Product: Year ended March 2013 media release*. Statistics New Zealand, Wellington. URL: http://www.stats.govt.nz/browse_for_stats/economic_indicators/NationalAccounts/RegionalGDP_MRYeMar13.aspx.

Stillman S, Maré DC (2009) The labour market adjustment of immigrants in New Zealand. *Economic Impacts of Immigration Working Paper Series*, The Department of Labour, Wellington

Warman C (2007) Ethnic enclaves and immigrant earnings growth (Enclaves ethniques et croissance des gains des immigrants). *The Canadian Journal of Economics / Revue canadienne d'Economique* 40: 401–422

Zhu P, Liu CY, Painter G (2014) Does residence in an ethnic community help immigrants in a recession? *Regional Science and Urban Economics* 47, 112-127