Longitudinal Analysis of Assimilation, Ethnic Capital and Immigrants' Earnings: Evidence from a Hausman-Taylor Estimation

Xingang (Singa) Wang[1], Sholeh Maani[2],

Paper prepared for New Zealand Association of Economists Conference 2012, Palmerston North, New Zealand, 27-29 June 2012.

Abstract

In this paper we examine the economic performance of immigrants using a panel/longitudinal approach. We extend the literature by incorporating the effect of 'ethnic capital' in our analysis in the Australasian setting. To construct social and resource networks for immigrant groups, we adopt the 'spatial model approach' to account for ethnic concentration and networks. We incorporate different measures of ethnic capital, in particular, ethnic group economic resources and spatial concentration. Moreover, we employ the Hausman-Taylor estimation method (1981) to account for potential endogeneity in the panel setting to examine the effects of ethnic capital and human capital using an eight-year Australian panel data set (HILDA). We find that immigrants tend to assimilate over time, but this effect is significantly affected by immigrants' ethnic group local concentration and resources. We further show that controlling for ethnic capital enhances the analytical explanation of the assimilation model.

Keywords: immigrant; assimilation; ethnic capital; ethnic concentration; ethnic enclave; panel/longitudinal; spatial model; Hausman-Taylor estimation; HILDA.

Acknowledgement: This paper uses unit record data from the Household, Income and Labour Dynamics in Australia (HILDA) Survey. The HILDA Project was initiated and is funded by the Australian Government Department of Families, Housing, Community Services and Indigenous Affairs (FaHCSIA) and is managed by the Melbourne Institute of Applied Economic and Social Research (MIAESR). The findings and views reported in this paper, however, are those of the authors and should not be attributed to AHURI, FaHCSIA or the MIAESR.

We also wish to thank Dr. Ping Yu for helpful comments.

[1] PhD Candidate, Economics Department, University of Auckland, New Zealand.
[2] Associate Professor, Economics Department, University of Auckland, New Zealand.

1. Introduction

"Assimilation" is an important indicator which measures the economic performance of immigrants. The word "assimilation" refers to the processes along which an immigrant's earning converges to a comparable native level, after residing in the host country for a certain period of time. As LaLonde and Topel (1991) pointed out, if new immigrants are not successfully assimilated, "increased immigrant flows may place additional burdens on public welfare systems, while exacerbating other social problems associated with persistent poverty". Therefore, the economic performance of immigrants is of special analytical and policy interest.

In 1978, Chiswick published his paper on the effect of Americanization on the earnings of foreign-born men. This paper laid a systematic theoretical foundation in the area of assimilation studies. Chiswick (1978) argued that in contrast with natives, immigrants are disadvantaged in the host country's labour market as they lack English language skills, social networks, knowledge of customs, information about job opportunities, and firm-specific training. Due to these weak points, new immigrants (especially those whose first language is not English) may face high barriers in finding a job. In addition, it might take a long time for their income to converge to the income level of natives in the host country.

However, economists hypothesise that immigrants will be assimilated eventually since immigrants continue to learn about the host country. Subsequent studies also observed significant influences on the assimilation process for immigrants from other factors: the quality of immigrant cohorts (Borjas, 1985), country of origin (e.g. Beenstock, Chiswick, & Paltiel, 2010; Borjas, 1987; Chiswick & Miller, 2008),

ethnic concentration (e.g. Edin, Fredriksson, & Aslund, 2003; Lazear, 1999) and personal English skill (e.g. Chiswick & Miller, 1995, 1996; Dustmann & Fabbri, 2003; McManus, Gould, & Welch, 1983).

A number of international studies have shown evidence of the assimilation process on immigrants around the world (e.g. Chiswick, 1978, 1980; Chiswick, Lee, & Miller, 2005; Constant & Massey, 2003; Fertig & Schurer, 2007). However, at the same time, other researchers could not confirm the assimilation process was significant and successful for all immigrant groups. For example, by testing synthetic cross-sectional data, Borjas (1985, 1995) found the assimilation effect was much weaker than had been reported in the previous cross-sectional studies in the United States. By examining the 1980, 1990 and 2000 US Census, Chiswick and Miller (2008) observed a strong "negative" assimilation effect on foreign-born men in the United States.

An increasing number of studies have paid attentions to the differences of assimilation effects across ethnic groups. It is also recognized that the assimilation processes of different ethnic groups have diverse patterns and time ranges. Borjas (1982) observed divergent assimilation processes for immigrants from Cuba, and Mexico to the United States. McDonald and Worswick (1999) have documented the persistence of income disparities between immigrants (from a non-English speaking background) and natives in Australia. Beenstock, Chiswick and Paltiel (2010) have found immigrants from Asia and Africa to Israel faced much greater earning disadvantages than those who migrated from the USSR; at the same time, European immigrants had higher incomes than natives in Israel.

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

Prior economic studies have provided little empirical evidence as to how ethnic factors influence immigrants' assimilation process and labour market performance. This paper uses "ethnic capital" as a key concept. In addition, we adopt the "spatial model" of network effects and a panel estimation approach and employ the Hausman-Taylor (HT) method (1981) to estimate the effects of ethnic capital and human capital on the relative earnings of immigrants.

This paper is arranged as follows: Section Two provides a brief description of "ethnic capital" and of certain hypotheses based on that concept. In Section Three we discuss the "spatial model" and the Hausman-Taylor Estimation approach adopted in this study. Section Four provides information on the data. We use an eight-year panel data set, the Household, Income and Labour Dynamics in Australia (HILDA) data, and the published 2001, 2006 Australian Census data. Empirical results and analyse are discussed in Section Five. The last section concludes this paper.

2. Immigrant Assimilation and "Ethnic Capital"

2.1 Ethnic capital

Borjas (1987) rooted the reasons for different assimilation profiles across ethnic groups in the effect of country of origin. He considered there are four factors which influence immigrants' labour market performance in the host country: age composition of immigrants, native language, political system and economic development of the source country.

The concept of "ethnic capital" was first put forward by Borjas (1992). He claimed that ethnicity plays a key role in the human capital accumulation process; and studied the effect of ethnic capital on skills in the immigrants' succeeding generation. The empirical evidence suggested that the skills of the immigrants' next generation significantly depend on both parental inputs and the quality of the ethnic environment (which Borjas calls "ethnic capital").

Borjas' theory incorporates the factors that stem from the country of origin. These factors are a kind of "innate" capital (and resources) of immigrants originating from their source country. This kind of capital cannot be easily altered by individual immigrants, since it is dependent on the overall macro-environment and the culture of the country of origin. Importantly, it belongs only to members of the same ethnic group and it cannot be utilised by others.

However, "innate" capital for immigrants is not from their country of origin only. For example, immigrants can access such capital from the host country as well, because at any time, earlier immigrants have already built up an ethnic environment (especially for social and commercial networks and other relative economic factors). Therefore, ethnic capital is a resource and also a capital which can be accessed by

subsequent immigrants from the same ethnic group. Such resource generated from the ethnic environment in the destination country is considered to have a more profound effect on immigrants' assimilation than the resources from their source country because they are created by previous cohorts of immigrants in the host country, and influenced by local socio-economic factors. In addition, this resource comes from immigrants themselves, so it can be adjusted and affected by immigrants. This also implies that the ethnic capital in the host country may vary over time, which is different from the nature of the "innate" capital from the country of origin.

This paper examines the effects of two "innate" capitals on immigrants' assimilation processes, and we refer to them in general as "ethnic capital".

We extend the definition of "ethnic capital" in this paper, by defining it as immigrant network based on: the country of origin; average skill level; group language proficiency; social network; geographical concentration; shared belief and other resources for a typical ethnic group. In other words, ethnic capital is the inherent trust and advantages which stem from, and belong to, a certain ethnic group. This is a new arena for immigration studies, particularly in the context of Australia, of which a comprehensive study is yet to emerge. This study is designed to enhance knowledge in this regard and therefore to fill this gap in the literature.

2.2 Hypotheses

To capture network effects, most previous international economic studies have adopted ethnic concentration/enclave as the proxy for networks for immigrants in the host country (e.g. Aguilera, 2009; Damm, 2009; Edin, et al., 2003; Toussaint-Comeau, 2008). Few other studies have used language group or language proficiency

(Bertrand, Luttmer, & Mullainathan, 2000; Chiswick & Miller, 2002). In this study, we construct a spatial network variable, endogenous "ethnic spatial lag" to represent the individual's network of economic resources in addition to ethnic concentration[3]. By doing so, we are able to separate the spatial network specific effect of the more general ethnic concentration/enclave. We hypothesize that both ethnic networks and ethnic concentration influence immigrants' economic performance.

(1) Ethnic network effect

Individuals are inherently linked through the groups they belong to. These groups include friendships, kinships, as well as race relations, or other relationships. Living in a common environment, produces shared experiences, knowledge, information and other products through these kinds of networks. Recent studies show that social networks may exert a significant influence on people's labour market performance (e.g. Frijters, Shields, & Price, 2005). For example, individuals may benefit from their friendships; their friends may introduce job opportunities to them, or provide them with assistance. Social networks are argued to be "the most profitable avenue of job search" for immigrants (Frijters, et al., 2005). For these reasons, individuals' labour market performance may not be independent and identically distributed (i.i.d.), especially for immigrants; thus, the labour market performance of an individual is correlated with that of other individuals to some extent. For these reasons, social networks may act positively on the process of immigrants' assimilation.

(2) Ethnic concentration

Recent international studies have generally indicated a negative effect of ethnic concentration on immigrants' earnings. For example, Chiswick and Miller (2002), and Bertrand et al. (2000) showed that linguistic concentration negatively

[3] Details are discussed in Section Three: Model Specifications.

influenced immigrants' labour market performance in the United States. In contrast, Edin et al. (2003) claimed that by correcting for the endogeneity of ethnic concentration, immigrants' earnings were positively correlated with the size of ethnic concentration in some cases in Sweden.

For example, immigrants' networks can affect their earnings through different channels. Immigrants may find greater opportunities of employment through geographic concentration. Firstly, an ethnic enclave creates job opportunities for immigrants by lowering the requirements for employment (e.g. having skill in the local language, or a recognised qualification). In addition, immigrant-owned businesses are considered to be the main source of employment opportunities for immigrants who come from the same ethnic group. It is observed that even after being located in the U.S. for 6 years, there were still around 40% of Cuban immigrants who were working for Cuban-owned businesses (Portes, 1987). Secondly, the immigrant market is potentially important for the local mainstream companies. Because native-born employees might know little about immigrants' culture and language; mainstream companies would like to hire immigrants to serve the target immigrant market.

Moreover, as discussed above, an ethnic enclave might increase the employment possibilities for immigrants in and out of that ethnic enclave. Therefore, immigrants may benefit from ethnic concentration, as more jobs could be generated by ethnic and geographic concentration. However, on the other hand, by lowering barriers to employment for immigrants, an ethnic enclave reduces the bargaining power of low-skilled immigrants, since it makes employment within the ethnic

enclave very attractive (e.g. working in an ethnic enclave can reduce the cost of learning English).

As a result, the effects of ethnic concentration on immigrants' assimilation might be different; for example, by ethnic group. Immigrants can be either "complements" or "substitutes" to each other. When the "substitute effect" is stronger than the "complement effect", immigrants compete for scarce employment opportunities in the host country labour market. Thus, under this kind of competition, immigrants may accept a lower salary than they would prefer in order to secure the employment opportunity. In the opposite case, if the "complement effect" dominates the "substitute effect" and with increasing in the proportion of immigrants in a specific region; a higher demand for immigrant labour would be generated, leading to more job opportunities and a higher salary for immigrants.

3. Model Specifications

In this section we discuss the modelling approach adopted. We incorporate a spatial component, and adjust for endogeneity in the panel setting through the Hausman-Taylor (1981) estimation method.

A cross-sectional model that is used to analyse how immigrant earnings respond to the assimilation process is:

$$lnW_i = c + \beta_1 X_i + \gamma_1 t_i + \gamma_2 {t_i}^2 + \beta_2 I_i + \varepsilon_i \quad (1)$$

Where W_i denotes the earnings of individual i in the host country; X is a vector of explanatory variables (for example, years of schooling completed, marital status, and years of labour market experience); t denotes years since migration to the host country; I is a dummy variable set to 1 if person i is foreign-born otherwise is 0; ε measures how earnings grow with the assimilation process (Borjas, 1985).

3.1 The spatial model

The spatial model expands the empirical framework to investigate the effect of ethnic capital. Under the ethnic capital hypothesis, individuals' incomes depend on ethnic capital and other socio-economic variables. Therefore, based on the basic econometric framework (3.1), we have controlled for the effects of cohort and ethnic capital, by incorporating ethnic concentration and a spatial weighted matrix effect of group economic resoruces as below:

$$y = \alpha l_n + \rho W y + \gamma Meth + \beta X + \varepsilon \quad (2)$$

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

Where W is a $n \times n$ ethnic spatial weight matrix which shows the first-order ethnic and geographical (ethnic-spatial) relationship among individuals. Wy is the endogenous "ethnic spatial lag", and it represents the linear combination of individual i's ethnic neighbours' labour market performances. *M* is the ethnicity matrix, and *eth* represents ethnic concentration. The parameter γ will reveal the effect of ethnic concentration. The vector X contains socio-economic variables and personal characteristics of individuals (e.g. education level, personal English proficiency level, years since migration, and immigrant identity). l_n is an identity vector of ones and associated with the parameters α and β. Thus, the coefficient ρ indicates the correlation of earnings among "ethnic neighbours" and also the size of the effect of the network in a specific region.

3.2 Ethnic spatial weight matrix

One can define individuals who are from the same ethnic group and location as the first-order "ethnic neighbours". Thus, "ethnic-spatial dependence" represents the case that an individual's labour market performance is influenced by its ethnic neighbours' labour market performance and other ethnic capital factors in that location.

Before the discussion of the ethnic-spatial relationship matrix *W*, the first-order ethnic-spatial neighbourhood matrix *E* will be introduced. Suppose P1, P2, P4 and P6 are all from Asia; P1 and P4 are all located in region A, while P2 and P6 located in region B. P3, P5 and P7 are from Europe, all of them located in region B. Thus, the 7×7 first-order ethnic-spatial neighbourhood matrix *E* is:

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

When the elements of the matrix *E* are zeros, individuals are not deemed to be the 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 an "ethnic spatial lag", matrix *E* should be normalised by unifying the row sums, and so we can form the ethnic spatial weight matrix *W*. For example:

$$W = \begin{array}{c} \\ P1 \\ P2 \\ P3 \\ P4 \\ P5 \\ P6 \\ P7 \end{array} \begin{array}{c} \begin{array}{ccccccc} P1 & P2 & P3 & P4 & P5 & P6 & P7 \end{array} \\ \begin{pmatrix} 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 1 \\ 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 1 \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1/2 & 0 & 1/2 & 0 & 0 \end{pmatrix} \end{array} \tag{4}$$

3.3Ethnic spatial autoregressive process

LeSage and Pace (2009) indicated the data generating process for the situation when the value of one observation *i* depends on the value of its neighbour *j*'s observation:

$$y_i = \alpha_i y_j + \beta X_i + \varepsilon_i \tag{5}$$

$$y_j = \alpha_j y_i + \beta X_j + \varepsilon_j \tag{6}$$

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

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

Thus, equations (5) and (6) imply a "simultaneous data generating process" that showing the dependence of y_i and y_j and vice versa. This feature leads us to a data generating process which is an "ethnic spatial auto regressive process" and we can have the following expression:

$$y_i = \rho \sum_{j=1}^{n} W_{ij} y_j + \varepsilon_i \qquad (7)$$

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

Where X_i is a vector of socio-economic variables for individual i. Since the "ethnic neighbour" is defined as individuals who are from the same ethnic group and settled in the same location; thus, $\sum_{j=1}^{n} W_{ij} y_j$ is the "ethnic spatial lag" in this case and it represents the linear combination of individual i's ethnic neighbours' labour market performances.

As a result, the matrix version of equation (7) is:

$$y = \rho W y + \varepsilon \qquad (8)$$

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

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.

3.4 Effect of network

Now, we can work out the model to investigate the effect of the network based on equation (2). Rearranging equation (2) result in:

$$(I_n - \rho W)y = \alpha l_n + \beta X + \varepsilon$$

$$y = (I_n - \rho W)^{-1}\alpha l_n + (I_n - \rho W)^{-1}\beta X + (I_n - \rho W)^{-1}\varepsilon \quad (9)$$

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

Since in previous economic models, immigrants' earnings estimation are based on a simplified specification as below:

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

We note that when we include the network effect, the "joint" effect of human capital ($(I_n - \rho W)^{-1}\beta$) in equation (9) where we have the "ethnic spatial lag" in our income regression model is larger than the normal coefficient for human capital β (in equation 3.9). In other words, without considering one of the effects of ethnic capital (the network effect), we may either underestimate or overestimate the effects of immigrants' personal characteristics and other socio-economic factors.

Furthermore, since $(I_n - \rho W)^{-1} = I_n + \rho W + \rho^2 W^2 + \rho^3 W^3 + \cdots$; so equation (9) can be extended into:

$$y = (I_n + \rho W + \rho^2 W^2 + \rho^3 W^3 + \cdots)\alpha l_n + (I_n + \rho W + \rho^2 W^2 + \rho^3 W^3 + \cdots)\beta X + (I_n + \rho W + \rho^2 W^2 + \rho^3 W^3 + \cdots)\varepsilon \quad (10)$$

As discussed before, W denotes the first-order ethnic-spatial relationship among individuals, and ρ shows the correlation with that individual's first-order

ethnic-spatial neighbours. W^2 represents the second-order ethnic spatial relationship; ρ^2 denotes the influence from that individual's second-order ethnic-spatial neighbours (that is neighbours' neighbours). Following the same logic; $(I_n - \rho W)^{-1}$ construct a full social networking for that individual and it capture all the information of network (e.g. Bonacich, 1972; Katz, 1953).

3.4 Effect of ethnic concentration

Immigrants' labour market performances are influenced by many ethnic capital factors, such as ethnic entrepreneurship, average language proficiency level, and ethnic concentration; and, in addition, the effects of those ethnic capital factors mentioned above differ across different regions under the hypotheses of ethnic capital. Thus, another ethnic spatial matrix *M* is needed when modelling the effects of other ethnic capital effects. This $m \times n$ matrix denotes the first-order relationship between an individual (person) and sub-ethnic groups; where m denotes the number of individuals in the data and n represents the number of sub-ethnic groups (when combining the ethnic group and locations together). For example, P1, P3, and P4 are Chinese, however, P1 is located in region 1, P3 and P4 are located in region 2; P2 and P5 are European, and P2 is located in region 1 while P5 is located in region 2. Therefore the ethnicity matrix *M* is:

$$M = \begin{array}{c} \\ P1 \\ P2 \\ P3 \\ P4 \\ P5 \end{array} \begin{pmatrix} \begin{array}{c} Chinese \\ Reg\,1 \end{array} & \begin{array}{c} Chinese \\ Reg2 \end{array} & \begin{array}{c} European \\ Reg1 \end{array} & \begin{array}{c} European \\ Reg2 \end{array} \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \tag{11}$$

Since every individual are specified belongs to only one sub-ethnic group the row sum would always equal 1 and the "row-stochastic" naturally holds for this matrix.

Therefore, the following model will explore effects of other ethnic capital factors: network and ethnic concentration effects.

$$y = \alpha l_n + \rho W y + \beta X + \gamma Meth + \varepsilon \quad (2)$$

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

Where *M* is the ethnicity matrix and *eth* represents ethnic concentration. The parameter γ represents the effect of ethnic concentration.

3.5 Hausman-Taylor (HT) Estimation (1981)

Hausman and Taylor (1981) have developed an econometric model for panel data, which allows controlling for endogeneity and at the same time the investigation of the effect of time-invariant variables. Their model characterises Fixed-Effect (FE), Random-Effect (RE), and OLS estimations. For example:

$$y_{it} = x1'_{it}\beta_1 + x2'_{it}\beta_2 + z1'_i\beta_3 + z2'_i\beta_4 + u_i + e_{it} \quad (12)$$

$$\text{Where, } E(e'_{it}|x1'_{it},\ x2'_{it},\ z1'_i,\ z2'_i) = 0,\ E(u_i|x1'_{it},\ z1'_i) = 0 \quad (13)$$

$$Cov(u_i, x1'_i) = 0,\ \text{ but } Cov(u_i|x2'_i) \neq 0$$

$$Cov(u_i, z13'_i) = 0,\ \text{ but } Cov(u_i|x4'_i) \neq 0$$

Therefore, this model contains time-variant variables, time-invariant variables and potentially endogenous variables. The Fixed-Effect (FE) model cannot estimate the coefficients β_3 and β_4 of time-invariant variables z1 and z2. At the same

time, the Random-Effect (RE) model is not efficient to estimate all of the effects since x2 and z2 are correlated with u.

However, the Hausman-Taylor (HT) estimation (1981) provides a solution. It follows three steps to estimate the coefficients[4]. The approach is equivalent to a three steps process:

(1) Estimate the above model by Fixed-Effect estimation with instruments ($x1_{it}$ and $z1_i$) to obtain $\widehat{\beta_{1_{FE-IV}}}$, $\widehat{\beta_{2_{FE-IV}}}$ and $\hat{\tilde{e}}_{it}$.

(2) Regress $\hat{\tilde{e}}_{it}$ on $z1_i$ and $z2_i$ with instrumental variables $x1_{it}$ and $z1_i$ by OLS in order to estimate $\widehat{\beta_{3_{IV}}}$, $\widehat{\beta_{4_{IV}}}$; then it allows us to calculate $\hat{\sigma}_e^2$ and the Random-Effects estimator θ_i.

(3) Derive the GLS transformation withθ_i and estimate the whole model by Random-Effect.

By investigating the bias and RMSE[5] properties of OLS, HT, FE, RE and pretest estimators from the Monte Carlo experiments, Baltagi, et al. (2003) argued that "OLS standard errors are biased and yield misleading inference under both a RE and HT world." When $\rho \neq 0$ (ρ is the "proportion of the total variance due to the individual effects"), HT, pretest and FE are the best estimation to estimate the coefficients of endogenous time-varying variables; HT is one of the best methods to estimate the coefficients of endogenous time-invariant variables. In addition, they found that with the increasing of individual effect, the bias for OLS and RE for the coefficients of the endogenous time-invariant variables and the corresponding RMSE

[4]More details are provided in the Appendix.
[5] Root-Mean-Square Error.

will also be increased. In addition, HT estimation overcomes disadvantages of traditional instrumental variable (IV) estimation. As Ruiz, Gomez, and Narvaez (2010) showed, efficient instruments need to satisfy three conditions: (1) rank condition, i.e. instruments must correlate with endogenous variables but not the error term; (2) exclusion restriction, i.e. instruments should be "legitimately excluded" from the earning function; (3) order condition, i.e. instruments need to be more numerous. However, without additional information provided by instruments; based on the nature of panel data, HT estimation can provide consistent and efficient measurement of the effects of assimilation and ethnic capital on immigrants in this case.

4. Data

4.1 The Household, Income and Labour Dynamics in Australia (HILDA)

The Household, Income and Labour Dynamics in Australia (HILDA) Survey is a household-based panel study which began in 2001. The wave 1 panel consisted of 7,682 households and 19,914 individuals. HILDA contains dynamic information about surveyed Australian natives' and immigrants' income, education, ethnicity, residence location, occupation, and family. In addition, HILDA divides Australia into 13 Major Statistical Regions. HILDA also provides fully detailed information about where immigrants come from.

A merged longitudinal data set is created based on data from the first eight waves of HILDA (from 2001 to 2008), and adopted in this study. In order to examine immigrants' labour market performance in Australia, only observations of full-time employed male immigrants and natives aged between 25[6] and 55 years have been employed. We use a balanced panel data set. Since some respondents refused to answer some questions, resulting in missing data, those individuals and the corresponding observations have been dropped from the data set. Because there are new, added and dropped respondents in each wave of the survey, longitudinal weights are applied in all regressions. As a result, the merged longitudinal data set contains 12,782 observations and 2,357 individuals; among them there are 517 immigrants, who contributed 2,662 observations.

We augment our data set by incorporating ethnic concentration Since HILDA collects information about the country of origin of individuals; it is possible to classify ethnic groups by parents' country of origin. However, the published 2001 and

[6]Students are excluded (e.g. Stillman & Maré, 2009).

2006 Australian Census data reports only information about individuals' country of birth. Therefore, in order to incorporate the Australian Census data with HILDA, the ethnicity of an individual will be classified by that individual's country of birth.

Immigrants from different ethnic backgrounds and countries of origin may have different assimilation processes. Thus, in order to examine the effect of ethnic capital on immigrants, they have been divided into five major groups, based on considerations of geography, and language. Asians and New Zealanders have been grouped as two single groups; while, due to the language effect, immigrants from the United Kingdom and Ireland have been categorized into one group, and immigrants from other European countries are all placed in a fourth group; the "rest of world" category contains all other immigrants.

An individual is categorized as being high-skilled if that person has obtained at least an advanced diploma or bachelor degree (e.g. Maani, 2004; Maani & Maloney, 2004). Since HILDA reports the age at which an individual left school, potential labour market experience is calculated by current age minus age of leaving school, as in other research (Gladden & Taber, 2002; Schultz, 1997). Wage has generally been considered as a good indicator of an individual's labour market performance by previous studies (e.g. Borjas, 1985); thus, in this paper real hourly wage will be examined. Real hourly wage is derived from HILDA by dividing weekly salary from an individual's main job by hours of work in that job. Furthermore, hourly wage has been adjusted by the Australian CPI[7].

[7]Base year is 1990.

4.2 2001 and 2006 Australian Census

To incorporate ethnic concentration information across Australia, we use the Australian Census for this data. We derived one of our two ethnic capital variables (ethnic concentration) from the published 2001 and 2006 Australian Census tables (Australian Bureau of Statistics, 2006, 2007). This ethnic capital variable is measured at the Australian Major Statistical Region (MSR) level[8]. We incorporated this data into HILDA data to examine immigrants' assimilation effect in Australia.

Ethnic concentration is defined in this study as:

$$\text{ethcon}_{ij} = \frac{\text{Population}_{ij}}{\text{Population}_{j}} \tag{14}$$

Where "i" denotes ethnic group (classified by country of origin), and "j" represents a specific region (at MSR level, totally 13 regions) in Australia. There are 51 countries of origin reported in the Census.

4.3 Demographic characteristics

Due to adjustments to Australian immigration policy during the past three decades, the structure of the immigrant population in Australia has been profoundly changed in relation to many aspects, such as country of origin, language skill, and education level. Therefore, in our analyses recent immigrants are examined as a separate group in order to show better the characteristics of recent and earlier immigrants. Recent immigrants are defined as immigrants who arrived in Australia after 1991.

[Table 1 Placed Here]

[8]Which is consistent with the location information reported by HILDA.

Table 1 represents the social-economic characteristics for full-time employed native and immigrant males aged between 25 and 55. The average age of recent immigrants is less than the average age of native males, while the average age of earlier immigrants is likely to be greater than that of both natives and recent male immigrants. It is noteworthy that half of the full-time employed recent male immigrants are high-skilled; this figure (53.61%) is higher than the corresponding figure for both natives (32.64%) and earlier immigrants (39.18%). However, earlier immigrants are more likely to be married; about 84.76% of them are married. Compared to earlier immigrants, recent immigrants arrived in Australia at an older age (29) than the earlier cohorts (16). Recent immigrants earned less than Australian native born workers with regard to their hourly wages. Most of immigrants are from "Main English Speaking Countries"[9], followed by Asian countries.

[9]It refers to U.K., Ireland, New Zealand, U.S.A., Canada, and South Africa.

5. Empirical Evidence

Recall that the main equation estimated in this paper examines the effects of ethnic capital, by incorporating ethnic concentration and a spatial weighted matrix effect of group characteristics as below:

$$y = \alpha l_n + \rho W y + \gamma Meth + \beta X + \varepsilon \quad (2)$$

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

Based on potential measurement error, selection bias, and other biases caused by un-observability (e.g. ability), some human capital variables (the skill level, English proficiency, and marital status) are treated as endogenous in our earnings models, as they have been in previous economic analyses (e.g. Card, 2000; Card, 1999; Chiswick & Miller, 1995, 1999; Garc á, Molina, & Navarro, 2008; Ruiz, et al., 2010). Moreover, due to neighbourhood effects and selection bias, the variable of ethnic concentration and ethnic network are also identified as endogenous (see Clark & Drinkwater, 2000; Edin, et al., 2003).

5.1 OLS Estimation

[Table 3 Placed Here]

Generally speaking, immigrant assimilation in Australia is confirmed by OLS estimations. The coefficients of "Years Since Migration" (YSM) in all models suggest that the hourly wage of immigrants is growing at a faster rate than that of natives by about 2% (when cohort effects have been controlled); at the same time, this rate is decreasing at a constant rate of around 0.02% (coefficients of YSM-square) per year's stay in Australia. This finding is consistent with other OLS studies in Australia

(e.g. Chiswick & Miller, 2002) and the United States (e.g. Borjas, 1995). When immigrants are pooled with natives, potential labour market experience increases wages for both natives and immigrants at a rate of 2% per year and this rate is also decreasing, by 0.04% annually. However, when we study this effect on immigrants only, the OLS estimation suggests a smaller effect of potential experience on immigrants' earnings than the general case and decreasing at a slower rate. Generally, married immigrants and natives tend to have a higher hourly wage than do unmarried individuals. Personal English skill and education level helps both male natives and immigrants to receive a higher hourly wage.

The network effects on immigrants' earnings assimilation are significant and their hourly earnings have a spatial correlation of approximately 0.007. Immigrants benefit from being spatially concentrated in Australia; that is, the coefficient of ethnic concentration is about 0.013 and it is statistically significant.

5.2 Hausman-Taylor Estimation

[Table 4 Placed Here]

Table 4 provides the regression results by Hausman-Taylor (HT) estimations. Overall, when we applied the HT estimation on HILDA data, some effects of endogenous variables are shown to be weaker than the OLS results; and the coefficients of exogenous variables are much stronger than the results from OLS estimations. In addition, now the coefficient of personal English skill is about 0.026 in the pooled sample case, and around 0.009 in the immigrants' case. The effect of skill level is also much weaker now in both the general and the immigrants' case. However, the first model suggests a much stronger initial earning disadvantage for immigrants.

In addition, all models suggest a stronger assimilation effect on immigrants, now the coefficients of YSM are around 0.04. This result is very close to Beenstock et al.,'s (2010) study on a panel model; and confirms that panel models suggest a much stronger effect of assimilation than do OLS models. As a result, the length of time they have been in the host country is likely to affect their search for sufficient information for local labour market and develop their social networks.

In addition, all models on cohort effects confirmed a significant improvement of quality of immigrants than the OLS estimation suggests. The HT estimations suggest stronger correlation in immigrants' hourly wage (now the coefficient is 0.008) than the OLS does. Compared to a weak significant positive effect of ethnic concentration on immigrants' hourly earnings (under OLS estimations), under the HT estimations, this effect becomes highly significant and larger (0.02).

Both of OLS and HT estimations suggest a positive and significant network effect on immigrants' earnings. This finding confirms the hypotheses about the effect of a network on immigrants' assimilation process: that is, their labour market performance is not independent and identically distributed; and their wages are correlated with each other and social networks act positively on immigrants' assimilation. However, the correlation of their hourly earnings is very low which suggests that they may need a stronger linkage and network to help their economic assimilation. Overall, immigrants benefit from spatial concentration, as such concentration is likely to result in more resources they can access once the ethnic population in a specific region is sufficiently large; therefore, when we take account

of the overall ethnic capital effects, ethnic capital acts positively on immigrants' hourly wage and confirms the hypotheses of ethnic capital.

Moreover, following the method of Ruiz et al., (2010), the Breusch-Pagan test (1980) has been applied on OLS residuals. The results suggest that the variance of individual effect α is not zero. In addition, from the HT estimations of ρ we can see that the unobservable individual error term is around 80% of the total error variance. Therefore, from the test results, one can conclude that the OLS estimator is not efficient. The HT estimator adopts the features of both a fixed-effect and random-effect model; and it provides the measurements of time-invariant variables as well as controls the endogeneity. Therefore, we think the HT estimation provides a better understanding of the effects of assimilation and ethnic capital on panel data.

5.3 Country of origin

[Table 5 Placed Here]

Table 5 summarises the specific effects of human capital by country groups. Results for immigrants from Asia, Major English Speaking Countries (ESC)[10], the United Kingdom, and New Zealand are considered individually. Since all immigrant respondents from ESC in our sample indicated they speak only English at home, so we treat them as proficient in English and therefore have dropped the dummy variable of "Proficiency in English" for them.

Generally speaking, the effects of YSM on immigrants from different countries are very similar, at approximately 0.05. However, the effects of potential

[10]Major English Speaking Countries (ESC) includes the United Kingdom, New Zealand, Canada, the United States, Ireland and South Africa.

experience are different. The effect of potential experience on immigrants from New Zealand is the highest (around 0.06) among all ethnic groups, followed by immigrants from Asia (0.04). The lowest return of potential experience appears among immigrants from the rest of world to Australia. Immigrants with good English skill significantly increase their hourly wage. For Asian immigrants, the coefficient is 0.03, and the coefficient for immigrants from the rest of world is 0.05. It seems that hourly earnings for immigrants from the rest of world start at a higher level than the hourly wage of other ethnic groups but increase at a lower rate (as the constant term and the coefficient skill level in the case of immigrants from rest of world are the highest among all immigrant ethnic groups, but the coefficient of YSM and potential experience for them are the lowest). Married Asian immigrants tend to obtain a higher hourly wage than other immigrants. Moreover, the HT estimation results showed sustained quality growth for immigrants in Australia across all immigrant ethnic groups.

Immigrants from the United Kingdom have the strongest network effect (0.07) among all immigrant groups. The network effect on Asian immigrants is also very strong (0.015).

The effects of ethnic concentration on immigrants with more mixed cultures and those different from Australian culture and language tend to be higher. For example, the coefficient of ethnic concentration for immigrants from rest of world are 0.04; for Asia, 0.03; for Major English Speaking Countries (ESC), -0.05; for the United Kingdom (UK), -0.01; and for New Zealand (NZ), -0.4.

For the effects of ethnic capital on immigrants' hourly earnings we find three interesting set of results: (1) The network effect is larger for immigrants from the

United Kingdom; (2) Immigrants from ESC (including the United Kingdom, New Zealand and other English-speaking countries) - a group of countries have similar language and culture background with Australia, the effects of ethnic concentration on these immigrants are negative and highly significant; (3) However, the ethnic concentration and network effects on Asian immigrants and immigrants from the rest of world are significantly positive and strong.

As a result, immigrants from the United Kingdom that speak the same language and share the same culture with Australia seem to have higher social and economic correlations; furthermore, they have a stronger network and are more economically linked. On the other hand, immigrants from ESC tend to be substitutes for each other due to the negative effect of ethnic concentration on their hourly earnings. Immigrants from Asia are significantly better off once they concentrate their location in a specific region in Australia. When more immigrants (from countries that speak a different language from and have a different culture to Australia) are located in the same region, they will "generate" demand for immigrant labour for themselves and off-set the initial disadvantages in the Australian labour market to some extent.

6. Conclusion

This paper examines the effect of ethnic capital, particularly ethnic network of economic resources and ethnic concentration on immigrants' assimilation. Empirical findings of OLS and HT estimations confirmed the effects of ethnic capital on immigrants' labour market performance in Australia. Ethnicity and language proficiency is further shown to be important for the assimilation process of immigrants.

We further find that the network variable plays a positive and significant effect on wage growth in all cases. A stronger social network and linkage help immigrants to achieve better economic performance and more successful assimilation. In addition, immigrants from the same cultural and language background as Australia (e.g. the United Kingdom) are more economically correlated compared to other immigrants.

Some recent international studies have observed negative effects of ethnic concentration (e.g. Bertrand, et al., 2000; Warman, 2007). Cutler and Glaeser (1997) in contrast have shown that African Americans in the United States received significantly higher wages once they lived in highly integrated areas. The empirical findings of this study show a similar pattern in Australia. We find that immigrant with a different cultural and language background from the Australian culture and language benefit from concentration and networking in a specific region in Australia. In addition, our study shows that when we controlled for both ethnic network effects and ethnic concentration, both factors have significant effects for all immigrant groups.

Finally, the results of this study strongly suggest greater attention should be given to the role of ethnic capital and immigrant networks on the assimilation process of immigrants.

7. References

Aguilera, M. B. (2009). Ethnic enclaves and the earnings of self-employed Latinos. *Small Buiness Economics, 33*, 413-425.

Australian Bureau of Statistics. (2006). 2001 Census tables. http://www.abs.gov.au/websitedbs/censushome.nsf/home/Data

Australian Bureau of Statistics. (2007). 2006 Census tables. http://www.abs.gov.au/websitedbs/censushome.nsf/home/Data

Beenstock, M., Chiswick, B. R., & Paltiel, A. (2010). Testing the immigrant assimilation hypothesis with longitudinal data. *Review of Economics of the Household., 8*(1), 7-27.

Bertrand, M., Luttmer, E. F. P., & Mullainathan, S. (2000). Network effects and welfare cultures. *The Quarterly Journal of Economics, 115*(3), 1019-1055.

Bonacich, P. (1972). Factoring and weighting approaches to status scores and clique identication. *Journal of Mathematical Sociology, 2*, 113-120.

Borjas, G. J. (1982). The Earnings of Male Hispanic Immigrants. *Industrial and Labor Relations Review, 37*, 222-239.

Borjas, G. J. (1985). Assimilation, Changes in Cohort Quality, and the Earnings of Immigrants. *Journal of Labor Economics, 3*(4), 463-489.

Borjas, G. J. (1987). Self-Selection and the Earnings of Immigrants. *The American Economic Review, 77*(4), 531-553.

Borjas, G. J. (1992). Ethnic capital and intergenerational mobility. *The Quarterly Journal of Economics, 107*(1), 123-150.

Borjas, G. J. (1995). Assimilation and changes in cohort quality revisited: What happened to immigrant earnings in the 1980s? *Journal of Labor Economics, 13*(2), 201-245.

Breusch, T., & Pagan, A. (1980). The lagrange multiplier test and its applications to model specification in econometrics. *Review of Economic Studies, 47*(1), 239-253.

Card, D. (2000). Estimating the returns to schooling: progress on some persistent econometric problems *NBER Working Paper 7769*. New York: National Bureau of Economic Research.

Card, D. (Ed.). (1999). *The causal effect of education on earnings* (Vol. 3A). Amsterdam: Elsevier.

Chiswick, B. R. (1978). The effect of Americanization on the earnings of foreign-born men. *The Journal of Political Economy, 86*(5), 897-921.

Chiswick, B. R. (1980). An analysis of the economic progress and impact of immigrants (pp. 410): The Employment and Training Administration, U.S. Department of Labor.

Chiswick, B. R., Lee, Y. L., & Miller, P. W. (2005). Immigrant earnings: A longitudinal analysis. *Review of Income and Wealth, 51*(4), 485-503. doi: 10.1111/1467-9361.00039

Chiswick, B. R., & Miller, P. W. (1995). The Endogeneity between Language and Earnings: International Analyses. *Journal of Labor Economics, 13*(2), 246-288.

Chiswick, B. R., & Miller, P. W. (1996). Ethnic Networks and Language Proficiency among Immigrants. *Journal of Population Economics, 9*(1), 19-35.

Chiswick, B. R., & Miller, P. W. (1999). Language skill and earnings among legalized aliens. *Journal of Population Economics, 12*(1), 63-89.

Chiswick, B. R., & Miller, P. W. (2002). Immigrant earnings: Language skills, linguistic concentrations and the business cycle. *Journal of Population Economics, 15*(1), 31-57.

Chiswick, B. R., & Miller, P. W. (2008). The "Negative" assimilation of immigrants: A special case (Discussion Paper No. 3563). *IZA Discussion Papers*. Bonn: Institute for the Study of Labor (IZA).

Clark, K., & Drinkwater, S. (2000). Enclaves, neighbourhood effects and employment outcomes: Ethnic minorities in England and Wales. *Journal of Population Economics, 15*(1), 5-29.

Constant, A., & Massey, D. S. (2003). Self-Selection, Earnings, and Out-Migration: A Longitudinal Study of Immigrants to Germany. *Journal of Population Economics, 16*(4), 631-653.

Cutler, D. M., & Glaeser, E. L. (1997). Are ghettos good or bad. *Quarterly Journal of Economics, 112*, 827–872.
Damm, A. (2009). Determinants of recent immigrants' location choices: quasi-experimental evidence. *Journal of Population Economics, 22*(1), 145-174.
Dustmann, C., & Fabbri, F. (2003). Language proficiency and labour market performance of immigrants in the UK. *The Economic Journal, 113*(489), 695-717.
Edin, P. A., 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.
Fertig, M., & Schurer, S. (2007). *Earnings Assimilation of Immigrants in Germany: The Importance of Heterogeneity and Attrition Bias*. The German Socio-Economic Panel (SOEP) Paper No. 30. German Institute for Economic Research. Berlin. Retrieved from http://www.diw.de/documents/publikationen/73/diw_01.c.61960.de/diw_sp0030.pdf
Frijters, P., Shields, M. A., & Price, S. W. (2005). Job Search Methods and Their Success: A Comparison of Immigrants and Natives in the UK. *The Economic Journal, 115*(507), F359-F376.
Garc á, I., Molina, J.-A., & Navarro, M. (2008). The effects of education on spouses'satisfaction in Europe. *Applied Economics, 42*(28), 3607-3618.
Gladden, T., & Taber, C. (Eds.). (2002). *Wage progression among less skilled workers*. New York: Russell Sage.
Hausman, J. A., & Taylor, W. E. (1981). Panel data and unobservable individual effects. *Econometrica, 49*, 1377–1398.
Katz, L. (1953). A new status index derived from sociometric analysis. *Psychometrika, 18*, 39-43.
LaLonde, R., & Topel, R. (1991). Immigrants in the american labour market: Quality, assimilation, and distribution effects. *American Economic Review, 81*(2), 297-302.
Lazear, E. P. (1999). Culture and language. *Journal of Political Economics, 107*, 95-126.
LeSage, J., & Pace, R. K. (2009). *Introduction to Spatial Econometrics (Statistics: A Series of Textbooks and Monographs)*: Chapman & Hall/CRC.
Maani, S. A. (2004). Why Have Maori Relative Income Levels Deteriorated Over Time? *The Economic Record, 80*(248), 100-123.
Maani, S. A., & Maloney, T. (2004). Returns to post-school qualifications: New evidence based on the HLFS Income Supplement (1997-2002), from http://www.dol.govt.nz/PDFs/PostSchoolQuals.pdf
McDonald, J. T., & Worswick, C. (1999). The Earnings of Immigrant Men in Australia: Assimilation, Cohort Effects, and Macroeconomic Conditions. *Economic Record, 75*(1), 49-62. doi: 10.1111/j.1475-4932.1999.tb02433.x
McManus, W., Gould, W., & Welch, F. (1983). Earnings of Hispanic men: The role of English language proficiency. *Journal of Labor Economics, 1*(2), 101-103.
Portes, A. (1987). The social origins of the Cuban enclave economy of Miami. *Sociological Perspectives, 30*(4), 340-372.
Ruiz, A. C., Gomez, L. N., & Narvaez, M. R. (2010). Endogenous wage determinants and returns to education in Spain. *International Journal of Manpower, 31*(4), 410-425.
Schultz, T. P. (Ed.). (1997). *Demand for children in low income countries* (1 ed. Vol. 1). Amsterdam: Elsevier B.V.
Stillman, S., & Mar é, D. C. (2009). *The labour market adjustment of immigrants in New Zealand*. Economic Impacts of Immigration Working Paper Series. The Department of Labour. Wellington. Retrieved from http://www.dol.govt.nz/publications/research/lmainz/lmainz.pdf
Toussaint-Comeau, M. (2008). Do ethnic enclaves and networks promote immigrant self-employment? *Economic Perspectives, 32*(4th Quarter), 30-50.
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*(2), 401-422.

8. Appendix

Table 1: Descriptive Statistics for Full-time Employed Male, Age 25-55, HILDA

	Australia-Born	Recent Immigrants	Earlier Immigrants
Age	39.3	37.8	43.3
High Skilled (%)	32.6	53.6	39.2
Married (%)	80.1	80.4	84.8
Age at First Arrival	-	29.1	16.4
Years Since Migration	-	8.6	26.9
Experience (potential)	22.8	20.6	26.4
Log of Real Hourly Wage in Main Job for High-Skilled*	2.9	2.8	2.9
Log of Real Hourly Wage in Main Job for Low-Skilled*	2.6	2.5	2.6
Born in Main English Speaking Countries (%)	-	41.3	57.7
Born in Rest of Europe	-	8.0	14.1
Born in Asia	-	34.0	20.1
Born in the Rest of World	-	16.8	8.1
Arrived between 2001 and 2008 (%)	-	9.7	-
Arrived between 1991 and 2000 (%)	-	90.3	-
Arrived between 1981 and 1990 (%)	-	-	44.6
Arrived between 1971 and 1980 (%)	-	-	24.0
Arrived before 1971 (%)	-	-	31.4
Number of Observations	10120	739	1923

Note: * All wages are adjusted by Australian CPI.

Table 2: Variable List and Definitions

Human Capital	
Potential Experience	This is a derived variable which is equal to an individual's current age minus the age of graduation.
Proficiency in English	This is a dummy variable, equal to one if that individual is proficient in English.
High Skilled	This is a dummy variable, equal to one if that individual obtained at least a Bachelor degree or Advanced Certificate.
Personal Characteristics	
Years Since Migration (YSM)	This variable represents the duration of immigration.
Married	This is a dummy variable, equal to one if that individual is married.
Arrived 2001-2008	This is a dummy variable, equal to one if that immigrant arrived between 2001 and 2008.
Arrived 1991-2000	This is a dummy variable, equal to one if that immigrant arrived between 1991 and 2000.
Arrived 1981-1990	This is a dummy variable, equal to one if that immigrant arrived between 1981 and 1990.
Arrived 1971-1980	This is a dummy variable, equal to one if that immigrant arrived between 1971 and 1980.
Arrived Before 1971	This is a dummy variable, equal to one if that immigrant arrived before 1971.
Ethnic Capital	
Network Effect	The average hourly wage of an individual's ethnic network.
Ethnic Concentration	The proportion of the population of a specific ethnic group to the total population size in a specific region.

Table 3: OLS Estimates of Log Hourly Wage: Full-time Employed Male Australian Natives and Immigrants, age 25-55, 2001-2008, HILDA

	Pooled Sample	Foreign-Born		
		(1)	(2)	(3)
Human Capital				
Potential Experience	0.0206***	0.0120**	0.0126**	0.0128**
	(0.00254)	(0.00575)	(0.00574)	(0.00574)
Potential Experience -squared	-0.000368***	-0.000126	-0.000140	-0.000143
	(0.0000542)	(0.000120)	(0.000120)	(0.000120)
Proficiency in English	0.339***	0.354***	0.344***	0.354***
	(0.0504)	(0.0560)	(0.0560)	(0.0563)
High Skilled	0.302***	0.294***	0.296***	0.301***
	(0.00850)	(0.0188)	(0.0187)	(0.0189)
Personal Characteristics				
Years Since Migration (YSM)	0.0174***	0.0168***	0.0162***	0.0155***
	(0.00442)	(0.00483)	(0.00482)	(0.00484)
YSM-squared	-0.000177**	-0.000198**	-0.000182**	-0.000171**
	(0.0000772)	(0.0000843)	(0.0000842)	(0.0000844)
Married	0.133***	0.0938***	0.0884***	0.0904***
	(0.00947)	(0.0224)	(0.0224)	(0.0224)
Immigrant	-0.422***	/	/	/
	(0.0898)	/	/	/
Arrived 2001-2008	0.441***	0.429***	0.443***	0.431***
	(0.0935)	(0.102)	(0.102)	(0.102)
Arrived 1991-2000	0.247***	0.229***	0.233***	0.237***
	(0.0680)	(0.0741)	(0.0739)	(0.0739)
Arrived 1981-1990	0.149***	0.130**	0.139**	0.143**
	(0.0524)	(0.0574)	(0.0573)	(0.0573)
Arrived 1971-1980	0.177***	0.171***	0.181***	0.185***
	(0.0393)	(0.0429)	(0.0429)	(0.0429)
Arrived Before 1971	Ref.	Ref.	Ref.	Ref.
Ethnic Capital				
Ln (Network Effect)	/	/	0.00856***	0.00666**
	/	/	(0.00239)	(0.00263)
Ln (Ethnic Concentration)	/	/	/	0.0134*
	/	/	/	(0.00768)
Constant	1.916***	1.597***	1.595***	1.644***
	(0.0578)	(0.131)	(0.131)	(0.134)
Observations	12782	2662	2662	2662
R-square	0.125	0.129	0.132	0.133
Breusch-Pagan Test (Chi-square)	1.59E+04	3.01E+03	2.99E+03	2.97E+03
Note: (1) Standard errors in parentheses				
(2) * p<0.10 ** p<0.05 *** p<0.01				
(3) For native-born, the age effect has been controlled.				

Table 4: Hausman-Taylor Estimates of Log Hourly Wage: Full-time Employed Male Australian Natives and Immigrants, age 25-55, 2001-2008, HILDA

	Pooled Sample	Foreign-Born			
		(1)	(2)	(3)	(4)
Human Capital					
Potential Experience	0.0453***	0.0426***	0.0421***	0.0426***	0.0422***
	(0.0000760)	(0.000187)	(0.000186)	(0.000187)	(0.000186)
Potential Experience -squared	-0.000637***	-0.000779***	-0.000769***	-0.000781***	-0.000771***
	(0.00000157)	(0.00000366)	(0.00000365)	(0.00000366)	(0.00000365)
Proficiency in English	0.0259***	0.00548***	0.00883***	0.00643***	0.00942***
	(0.00153)	(0.00168)	(0.00168)	(0.00168)	(0.00168)
High Skilled	0.138***	0.0424***	0.0428***	0.0454***	0.0449***
	(0.000842)	(0.00210)	(0.00209)	(0.00210)	(0.00209)
Personal Characteristics					
Years Since Migration (YSM)	0.0426***	0.0536***	0.0541***	0.0530***	0.0537***
	(0.000105)	(0.000135)	(0.000135)	(0.000136)	(0.000135)
YSM-squared	-0.000491***	-0.000486***	-0.000491***	-0.000479***	-0.000486***
	(0.00000208)	(0.00000239)	(0.00000238)	(0.00000239)	(0.00000239)
Married	0.0501***	0.177***	0.177***	0.178***	0.177***
	(0.000277)	(0.000667)	(0.000666)	(0.000667)	(0.000666)
Immigrant	-0.999***	/	/	/	/
	(0.00246)	/	/	/	/
Arrived 2001-2008	1.093***	1.465***	1.496***	1.462***	1.493***
	(0.00401)	(0.00508)	(0.00505)	(0.00508)	(0.00505)
Arrived 1991-2000	0.703***	1.026***	1.039***	1.043***	1.052***
	(0.00233)	(0.00307)	(0.00306)	(0.00309)	(0.00308)
Arrived 1981-1990	0.383***	0.620***	0.631***	0.635***	0.642***
	(0.00207)	(0.00266)	(0.00264)	(0.00267)	(0.00266)
Arrived 1971-1980	0.265***	0.418***	0.425***	0.429***	0.433***
	(0.00201)	(0.00247)	(0.00245)	(0.00248)	(0.00246)
Arrived Before 1971	Ref.	Ref.	Ref.	Ref.	Ref.
Ethnic Capital					
Ln (Network Effect)	/	/	0.00861***	/	0.00839***
	/	/	(0.0000626)	/	(0.0000629)
Ln (Ethnic Concentration)	/	/	/	0.0288***	0.0208***
	/	/	/	(0.000589)	(0.000591)
Constant	1.922***	0.608***	0.586***	0.724***	0.670***
	(0.00180)	(0.00366)	(0.00365)	(0.00436)	(0.00437)
Observations	12782	2662	2662	2663	2662
sigma_u	0.5446	0.6780	0.6720	0.6776	0.6724
sigma_e	0.2579	0.2880	0.2874	0.2879	0.2874
rho	0.8168	0.8472	0.8454	0.8471	0.8456
Wald Chi-square	1.47E+06	6.90E+05	7.14E+05	6.92E+05	7.15E+05
Note: (1) Standard errors in parentheses					
(2) * p<0.10 ** p<0.05 *** p<0.01					
(3) For native-born, the age effect has been controlled.					

Table 5: Hausman-Taylor Estimates of the Effects of Ethnic Capital on Log Hourly Wage: Full-time Employed Male Immigrants in Australia, age 25-55, 2001-2008, HILDA

	Asia	Rest of Word	ESC^	UK	New Zealand
Human Capital					
Potential Experience	0.0435***	0.0140***	0.0394***	0.0260***	0.0570***
	(0.000396)	(0.000334)	(0.000271)	(0.000365)	(0.000556)
Potential Experience -squared	-0.00112***	-0.000144***	-0.000596***	-0.000331***	-0.000986***
	(0.00000827)	(0.00000678)	(0.00000506)	(0.00000646)	(0.0000111)
Proficiency in English	0.0262***	0.0479***	/	/	/
	(0.00214)	(0.00371)	/	/	/
High Skilled	0.154***	0.357***	0.00871***	0.000990	0.138***
	(0.00717)	(0.00583)	(0.00211)	(0.00224)	(0.00532)
Personal Characteristics					
Years Since Migration (YSM)	0.0541***	0.0379***	0.0550***	0.0559***	0.0542***
	(0.000290)	(0.000258)	(0.000190)	(0.000290)	(0.000385)
YSM-squared	-0.000221***	-0.000171***	-0.000626***	-0.000655***	-0.000829***
	(0.00000633)	(0.00000452)	(0.00000311)	(0.00000409)	(0.00000812)
Married	0.468***	0.0920***	0.0230***	0.00614***	0.0637***
	(0.00153)	(0.00129)	(0.000865)	(0.00122)	(0.00135)
Arrived 2001-2008	1.434***	1.436***	1.471***	1.641***	1.306***
	(0.0119)	(0.00954)	(0.00779)	(0.0130)	(0.0162)
Arrived 1991-2000	1.221***	1.033***	0.928***	0.774***	0.743***
	(0.00852)	(0.00666)	(0.00438)	(0.00701)	(0.0122)
Arrived 1981-1990	0.741***	0.937***	0.473***	0.505***	0.386***
	(0.00757)	(0.00597)	(0.00321)	(0.00437)	(0.0111)
Arrived 1971-1980	0.627***	0.460***	0.374***	0.466***	0.350***
	(0.00696)	(0.00503)	(0.00306)	(0.00407)	(0.0108)
Arrived Before 1971	Ref.	Ref.	Ref.	Ref.	Ref.
Ethnic Capital					
Ln (Network Effect)	0.0148***	0.00583***	0.00722***	0.0721***	0.00830***
	(0.000133)	(0.0000831)	(0.000117)	(0.00142)	(0.000177)
Ln (Ethnic Concentration)	0.0303***	0.0358***	-0.0467***	-0.0112***	-0.367***
	(0.000842)	(0.00125)	(0.00156)	(0.00255)	(0.00319)
Constant	0.328***	0.886***	0.718***	0.833***	-0.719***
	(0.0102)	(0.0108)	(0.00605)	(0.00916)	(0.0162)
Observations	638	609	1415	826	381
sigma_u	0.7230	0.6146	0.6302	0.7468	0.7378
sigma_e	0.3379	0.2535	0.2590	0.2486	0.2812
rho	0.8208	0.8546	0.8555	0.9002	0.8731
Wald Chi-square	3.38E+05	1.56E+05	3.29E+05	1.37E+05	1.30E+05
Note: (1) Standard errors in parentheses					
(2) * p<0.10 ** p<0.05 *** p<0.01					
(3) ESC^ stands for Major English Speaking Countries:					
United Kingdom, New Zealand, Canada, USA, Ireland and South Africa.					

Hausman and Taylor estimation (1981)

Assume the under panel model:

$$y_{it} = x1'_{it}\beta_1 + x2'_{it}\beta_2 + z1'_i\beta_3 + z2'_i\beta_4 + u_i + e_{it}$$

Where, $E(e'_{it}|x1'_{it},\ x2'_{it},\ z1'_i,\ z2'_i) = 0,\ \ E(u_i|x1'_{it},\ z1'_i) = 0$

$$Cov(u_i, x1'_i) = 0,\ \ but\ \ Cov(u_i|x2'_i) \neq 0$$

$$Cov(u_i, z1'_i) = 0,\ \ but\ \ Cov(u_i|z2'_i) \neq 0$$

Therefore, this model contains time-variants variables, time-invariant variables and endogenous variables. The Fixed-Effect model cannot estimate the coefficients β_3 and β_4 of time-invariant variables$z1$and $z2$. At the same time, Random-Effect model is not efficient to estimate all of the effects since x2 and z2 are correlated with u.

However, Hausman and Taylor estimation (1981) provides a solution and it follows three steps to estimate the coefficients:

(1) Estimate the above model by Fixed-Effect estimation with instruments ($x1_{it}$ and $z1_i$) to obtain $\widehat{\beta_1}_{FE-IV}$, $\widehat{\beta_2}_{FE-IV}$ and $\hat{\tilde{e}}_{it}$:

$$\because\ y_{it} = x1'_{it}\beta_1 + x2'_{it}\beta_2 + z1'_i\beta_3 + z2'_i\beta_4 + u_i + e_{it}$$

$$\text{and}\ \ \bar{y}_{it} = \overline{x1}'_{it}\beta_1 + \overline{x2}'_{it}\beta_2 + z1'_i\beta_3 + z2'_i\beta_4 + u_i + \bar{e}_{it}$$

$$\therefore y_{it} - \bar{y}_{it} = (x1'_{it} - \overline{x1}'_{it})\beta_1 + (x2'_{it} - \overline{x2}'_{it})\beta_2 + (e_{it} - \bar{e}_{it})$$

$$\therefore\ \ \tilde{y}_{it} = \widetilde{x1}'_{it}\beta_1 + \widetilde{x2}'_{it}\beta_2 + \tilde{e}_{it}$$

By Fixed-Effect estimation with instruments ($x1_{it}$ and $z1_i$), we can have:

$$\hat{\tilde{e}}_{it} = \tilde{y}_{it} - \widetilde{x1}'_{it}\beta_{1_{FE-IV}} - \widetilde{x2}'_{it}\beta_{2_{FE-IV}}$$

(2) Regress $\hat{\tilde{e}}_{it}$ on $z1_i$ and $z2_i$ with instrument variables $x1_{it}$ and $z1_i$ by OLS in order to estimate $\widehat{\beta_{3_{IV}}}$, $\widehat{\beta_{4_{IV}}}$; then it allows us to calculate $\hat{\sigma}_e^2$ and the Random-Effects estimator θ_i:

Applying OLS to estimate $\hat{\tilde{e}}_{it} = z1_i'\beta_3 + z2_i'\beta_4 + v_i$ with instruments $x1_{it}$ and $z1_i$.

Define $\tilde{\varepsilon}_{it} = \tilde{y}_{it} - \widetilde{x1}_{it}'\widehat{\beta_{1_{FE}}} - \widetilde{x2}_{it}'\widehat{\beta_{2_{FE}}} - z1_i'\widehat{\beta_{3_{IV}}} - z2_i'\widehat{\beta_{4_{IV}}}$

Then $\hat{\sigma}_e^2 = \hat{\varepsilon}'\hat{\varepsilon}/_{NT} => \theta_i = 1 - \sqrt{\sigma_e^2/(\sigma_e^2 + T_i\sigma_u^2)}$

(3) Deriving the GLS transformation with θ_i and estimate the whole model by Random-Effect.

$\because y_{it} = x1_{it}'\beta_1 + x2_{it}'\beta_2 + z1_i'\beta_3 + z2_i'\beta_4 + u_i + e_{it}$

$$\therefore y_{it} - \theta_i\bar{y}_{it} = (x1_{it}' - \theta_i\overline{x1}_{it}')\beta_1 + (x2_{it}' - \theta_i\overline{x2}_{it}')\beta_2 + (1-\theta_i)z1_i'\beta_3 + (1-\theta_i)z2_i'\beta_4 + [(1-\theta_i)u_i + (e_{it} - \theta_i\bar{e}_{it})]$$

$$\therefore \tilde{y}_{it} = \widetilde{x1}_{it}'\beta_1 + \widetilde{x2}_{it}'\beta_2 + \widetilde{z1}_i'\beta_3 + \widetilde{z2}_i'\beta_4 + \tilde{\varepsilon}_{it}$$