The immigrant wage gap and assimilation in Australia: does unobserved heterogeneity matter?

Robert Breunig[1], Syed Hasan and Mosfequs Salehin

Australian National University

31 July 2013

Abstract

Immigrants to Australia are selected on observable characteristics. They may also differ from natives on unobservable characteristics such as ambition or motivation. If we account for unobservable differences, we find a wage gap for immigrant men from English speaking backgrounds, in contrast with previous research which has found no wage gap. Controlling for unobserved heterogeneity also seems important for finding cohort effects. Immigrants that arrived before 1985 faced a larger wage gap compared to native-born Australians than subsequent cohorts. Confirming other research, we find wage gaps for immigrant men and women from non-English speaking backgrounds. Wage assimilation occurs slowly for all groups, but is slowest for those from non-English speaking backgrounds.

JEL Codes: J31, J61

Keywords : immigrants; wage gap; assimilation; Australia; cohort effects

[1] Corresponding author: Robert.Breunig@anu.edu.au, Crawford School of Public Policy, Australian National University, Canberra ACT 0200, Australia. We would like to thank Juergen Meinecke, Mathias Sinning and participants at the 2012 Australian Conference of Economists and the 2012 Labour Econometrics Workshop for helpful comments. 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 (Melbourne Institute). The findings and views reported in this paper, however, are those of the authors and should not be attributed to FaHCSIA or the Melbourne Institute or the Federation of Associated Organizations.

1 Introduction

In this paper, we re-examine the wage gap on entry and wage assimilation of immigrants to Australia. Australia is a land of immigrants, with over 25% of the population born overseas and net population growth heavily driven by migration. In 2010 Australia had the third highest proportion of overseas born residents in the world (ABS, 2011). The relationship between immigrant and non-immigrant wages is one important aspect of understanding the immigrant experience.

Although a number of Australian studies have explored the wage assimilation of immigrants, few have used panel data to account for unobservable differences between migrants and non-migrants. Immigrants may differ from native-born Australians in unobservable characteristics. These unobservable effects, if not accounted for, can lead to an omitted variable problem, potentially biasing estimates.

We use the Hausman and Taylor (1981) estimator which allows estimation of time-invariant variables, e.g. immigrant status, whilst controlling for unobserved individual heterogeneity. Using panel data from the Household Income and Labour Dynamics in Australia (HILDA) survey, we find a wage gap for immigrant men from English speaking backgrounds in contrast to previous research. Wages for men from both English and non-English speaking backgrounds approach those of native Australians over time, but at a much faster rate for the former group. For working women we find no significant differences in the initial wage gap nor in the assimilation profile when we control for unobserved heterogeneity. This may be due to selection, as discussed below. Finally, when we control for unobserved heterogeneity and split immigrants into separate arrival cohorts, wage gaps and assimilation profiles differ by cohort. Previous literature which does not account for unobserved heterogeneity fails to find cohort effects. These effects may be interpreted as changes in cohort quality over time and they may be due to changes in Australian immigration policy and economic conditions affecting the selectivity of immigrants to Australia.

2 Background, assimilation and unobserved heterogeneity

Most Australian studies have utilised cross-sectional data and the standard human capital function modified for immigrant adjustment as used by Chiswick (1978).

Chiswick and Miller (1985) use 1981 Australian Census data and find that male immigrants have seven per cent lower incomes than comparable native born men, but find no earnings disadvantage for second generation migrants. They find that immigrants have lower returns on their home country education and work experience than native born men, particularly immigrants from non-English speaking countries. In general most studies find that migrants from non-English speaking countries earn less than their Australian counterparts, but those from English speaking countries have similar outcomes to the native-born (Preston, 2001, p108). A few studies have even found that migrants from some English speaking countries earn more than comparable Australian-born workers (Chapman and Mulvey, 1986, Langford, 1995).

Other cross-sectional research has sought to explain immigrant–native born earnings differentials. International transferability of human capital (Chiswick and Miller, 2010, Chapman and Iredale, 1993, Beggs and Chapman, 1991), English language fluency (Chiswick and Miller, 1995), labour market conditions in Australia at the time of migration (McDonald and Worswick, 1999) and age at migration (Wilkins, 2003) all help to explain the wage gap. Chiswick and Miller (1985) find that immigrants' income increases with duration of residence, but McDonald and Worswick (1999) find little or no earnings assimilation for immigrants from non-English speaking countries. Second generation migrants, which we do not consider, have also received considerable attention in the Australian literature. Generally, authors find no wage gap (Chiswick and Miller, 1985) nor any wealth disadvantage (Doiron and Guttmann, 2009). One exception is Messinis (2009) who uses a cross-section of data from HILDA and finds that second generation migrants from English speaking backgrounds earn less than comparable Australians.

Using cross-sectional data may create biased estimates of the assimilation process if there is selective out-migration, or if individuals arriving at different points in time differ in unobserved human capital characteristics. Borjas (1989) and Lubotsky (2007) both find that selective out-migration of lower quality immigrants has overstated the wage progress and assimilation of immigrants to the United States in previous studies using cross-sectional data. Selective out-migration does not seem to be a major issue in Australia which is a final destination country for most immigrants.[2] Unobserved heterogeneity, as pointed out by Cobb-Clark (2003), is probably the more important issue, 'as changes in the state of the Australian labour market and the generosity of Australian income support policy would have directly affected returns to migration, altering the selectivity of the immigrant stream.' This paper takes up the challenge of attempting to address this issue.

Borjas (1985) demonstrates that unobserved heterogeneity among immigrant cohorts can bias estimates of years since migration on relative wage outcomes. The effect of years since migration on earnings is biased upwards if new immigrants are more able than previous arrivals. In Australia, the evidence is mixed. Beggs and Chapman (1988) find evidence of cohort effects for immigrants from non-English speaking backgrounds. However, McDonald and Worswick (1999) find no evidence that unobserved cohort quality of immigrants has changed over time.

Panel data can be used to control for unobserved individual heterogeneity (Baltagi, 2005) and selective out-migration (Borjas, 1989; Lubotsky, 2007). However, in the presence of period effects, it can also lead to a spurious finding of no assimilation; see Beenstock et al. (2010). To our knowledge, two Australian studies use longitudinal data. Chiswick et al. (2005), using the Longitudinal Survey of Immigrants to Australia, find wage equation estimates similar to cross-sectional ones. The data does not allow them to calculate wage gaps or assimilation profiles.

Cobb-Clark et al. (2012) use the HILDA data to compare alternative methods for estimating immigrant wage and employment assimilation. They conclude that fixed effects estimates on an unbalanced panel are preferable but they find no significant differences in wage assimilation profiles across alternative methods. Their fixed effects estimates are similar to ours but they do not employ the Hausman-Taylor estimator, which is our preferred specification.

Several international studies have used panel data to control for unobserved individual heterogeneity. Hum and Simpson (2004) estimate immigrant earnings and find that immigrant earnings assimilation in Canada is slower than previously thought.

[2] According to table 6.1 in Commonwealth of Australia (2013), skilled migrants make up about 50 per cent of entrants over the past few years. Of those who emigrate, skilled migrants make up about 60 per cent. Furthermore, out-migration is less than 10 per cent of new immigrants. Given this small difference in skill mix and the dis-proportionate size of the new arrival group, it seems that skill mix is a minor issue.

Fertig and Schurer (2007) find a similar result using German panel data. For the US, Hu (2000) finds little or no immigrant assimilation once longitudinal data is used. Lubotsky (2007) shows that studies utilising repeated cross-sections or synthetic cohorts in the US have overstated the assimilation and wage growth of immigrants. Using longitudinal earnings records from 1951 to 1997 for the US, he found that immigrant earnings growth was considerably slower than had been predicted using repeated cross-sections. [3]

Against this background, the contribution of our paper is to study the wage assimilation experience of Australian immigrants using panel data and accounting for the role of unobserved heterogeneity. In the next section we discuss our data and then proceed to our empirical approach and results.

3 Data

The data used is from the first nine waves (2001–2009) of the Household, Income and Labour Dynamics in Australia (HILDA) survey. Wooden and Watson (2007) provide a detailed overview of HILDA. HILDA is a nationally representative longitudinal survey of Australian households. It began in 2001 and approximately 7,000 households and 13,000 individuals have responded in every wave.

We use HILDA data to create two analysis sub-samples. For the first sample, we pool the observations over all nine waves to create a pooled cross-section. The pooled cross-section sample is used to estimate a baseline model for comparison with previous cross-sectional studies in Australia. It also provides a comparison with our panel data estimates. The second sample uses the HILDA data as an unbalanced panel over nine waves to estimate fixed effects, random effects and Hausman-Taylor panel data estimators. In both the panel and pooled cross-section, we consider men and women separately.

Our sample is restricted to men and women aged between 24 and 59 years of age, to exclude those facing decisions about full-time study or retirement. In addition, full-time students are excluded even if they reported being employed. We also exclude

[3] Cobb-Clark and Connolly (1997), Antecol et al. (2003), Richardson and Lester (2004) and Antecol et al. (2006) discuss similarities and differences of the Australian, U.S., and Canadian experiences and policy frameworks.

individuals who are self-employed or working in a family business. Individuals who refused to disclose their country of origin or their year of arrival to Australia or those who report working positive hours but have missing or zero hourly wages are excluded. Those who reported working more than 60 hours or less than 5 hours a week are also excluded to minimise measurement error in the hourly wage. Those with missing or incomplete work experience information are excluded[4]. Finally, we exclude individuals who are retired or have stopped working due to illness, injury or disability. The exclusions listed above are common to both analysis sub-samples.

For the panel sample we also exclude all individuals who are not employed. A small number of individuals[5] in our panel sample acquired greater amounts of education with time. In these cases we assign an education level to them based on an average of their education level during the panel thus treating education as time-invariant. The number of observations by wave is reported in Table 1. Means and standard deviations of key variables for wave 5 of the panel sample are provided in Table 2. Years since migration is defined using the HILDA variable which is based upon the interviewer asking "When was the first time you came to Australia to live for 6 months or more (even if you have spent time abroad since)?"

[Table 1 about here]

For the pooled cross-section, we drop observations if the partner has incomplete wage or employment information, or if the partner is self-employed. We make these exclusions, only for the pooled cross-section, because we estimate a sample selection model for employment that uses partner information. Those with missing work experience information are also dropped. Sample statistics are similar to the panel sample in Table 2 and are available upon request from the authors[6].

[Table 2 about here]

[4] Our substantive results are not influenced by this exclusion based on working hours which affects less than one per cent of the sample. Exclusion of those with missing work experience information reduces the sample by less than one per cent. The reported wage assimilation profiles below do not change if we include these individuals in the sample and control for the missing experience values with a dummy variable.
[5] These individuals comprised less than 5% of the sample of men and less than 7% of the sample of women.
[6] Throughout the paper we discuss results and sensitivity testing which space constraints prevented us from presenting. All of these are available upon request from the authors even where not explicitly stated.

Hourly wage is defined as the gross weekly salary of the individual from all jobs divided by the total number of hours worked in that week. Immigrant is a dummy variable which is equal to 1 if the individual is born outside of Australia. English speaking background (ESB) is equal to 1 if an immigrant is from the United Kingdom (UK), New Zealand, Canada, USA, Ireland or South Africa; all other immigrants are defined has having a non-English speaking background (NESB). Indicator dummies for education are mutually exclusive and indicate the highest level of education achieved. Partnered status includes both marriages and de-facto relationships.

Immigrants comprise 22% of our sample, less than the official estimate of 25% (ABS, 2009). This is due to under-representation of immigrants (see above) and partly to the age exclusions we impose. In the panel estimates, we only consider employed individuals which may also have an effect on the percentage of immigrants in our analysis sample. Both male and female immigrants earn more on average than their Australian counterparts; see Table 2. This is not surprising since immigrants in the sample are better educated, older, have greater work experience and mainly stay in urban areas. Native-born Australians, on the other hand, are more likely to be in paid employment than immigrants[7]; this is especially so for women.

One concern related to using HILDA is that while re-interview rates for immigrants born in mainly English speaking countries are quite similar to those for survey participants born in Australia, they are about 10 per cent lower in wave 1 and about 13 per cent lower across the first five waves of the survey for immigrants born in non-English speaking countries (Wooden and Watson, 2007)[8].

In order to better understand the possible effect of attrition in our data, we compare education levels between those who remain in the survey and those who drop out of the survey. Table A1 in the appendix compares the education levels (more than high school education compared to only high school or less) for three different groups—ESB and NESB immigrants and non-immigrants. The tables are constructed by

[7] For men, the employment rate for Australian-born is 90.6% in the pooled cross-sectional sample as compared to 88.6% for immigrant men. For women, these rates are 76.0% and 72.3% respectively.

[8] Attrition in HILDA is low compared to other major panel surveys. Attrition from wave 1 to wave 2 was over 10 per cent but only 3.7 per cent from wave 8 to wave 9. 71.5 per cent of all wave 1 respondents were present in wave 9. For the sub-sample of those born in non-English speaking countries, 62.3 per cent of wave 1 respondents are still present at wave 9. For details of the attrition analysis and general information on the survey, see Melbourne Institute of Applied Economic and Social Research (2010).

looking at education level for those present in the panel at time *t* split by whether or not they are still in the sample at time *t+1*. For this table, we pool across all waves. Wave-by-wave analysis is strikingly similar.

For all three groups, we see that more educated people are more likely to remain in the survey. This effect is strongest for NESB—73 per cent who are in the sample at both time *t* and *t+1* have greater than high school education whereas only 66 per cent of those who were in the sample at time *t* but not at time *t+1* have greater than high school education. The difference is four per cent for non-immigrants and only one per cent for ESB immigrants. These differences in education level for panel leavers and stayers seem fairly small and are thus likely to have only small implications for our results. We return to this in the concluding section.

HILDA does not include information on entry visa type for immigrants preventing us from distinguishing between immigrants who arrive through the skilled migrant program and those who arrive through other programs such as family reunification for which selectivity is not related to employability. We also return to this issue below.

4 Empirical Strategy

4.1 Hausman Taylor Estimator

Hausman and Taylor (1981) - hereafter HT - formulated an instrumental variable estimator for panel data that controls for possible correlation between included variables and unobserved individual effects. The standard fixed effects estimator can control for unobservable individual effects but it does not allow estimation of any included time-invariant variables. The HT estimator, described in detail in Breusch et al. (1989) (referred to as BMS hereafter), allows estimation of included time-invariant variables, provided that the number of included exogenous variables that are varying over both individuals and time are greater than the number of included endogenous variables that are time invariant. External instruments are not required; instruments are derived from within the model. The HT estimator may improve efficiency relative to standard fixed effects. The order condition for the existence of the HT estimator is that the number of included time-varying variables that are assumed to be uncorrelated with the unobservable individual effects has to be greater than or equal to the included time-invariant variables that are correlated with the individual effects.

Amemiya and MaCurdy (1986) - hereafter AM - and BMS both propose estimators that are more efficient than the HT estimator but which impose stronger exogeneity assumptions, see Baltagi (2005, p127). Cornwell and Rupert (1988) confirm that AM and BMS estimators are more efficient than the HT estimator in their analysis of returns to schooling, but their results are disputed by Baltagi and Khanti-Akom (1990). The AM and HT estimators have been found to produce similar estimates (e.g., Hum and Simpson, 2004). Implementing AM and BMS in unbalanced panels requires additional assumptions to deal with missing observations and individual spells which do not start at the same time period. AM and BMS also impose stronger exogeneity assumptions than the HT estimator. Our unbalanced panel and the weaker exogeneity assumptions required motivate our choice of the HT estimator.

4.2 Panel Model Specification

We estimate three wage equations each for men and women using the natural log of hourly wage as dependent variable. Table 3 lists the variables used in the wage equations. Age is only included as a quadratic because its level is perfectly co-linear with the wave dummies. The second equation allows ESB and NESB migrants to have different assimilation profiles. Lastly, we estimate a wage equation that has dummies for different arrival cohorts of immigrants. We estimate wage equations for men and women separately, as returns to human capital and labour market outcomes generally vary between men and women (Preston (2001), p102).

[Table 3 about here]

Many authors have estimated separate wage equations for the native-born and immigrants to allow for differing rates of return to characteristics (Beggs and Chapman, 1988; Chiswick and Miller, 1985). We test this by estimating a random effects model with interaction terms between included variables and immigrant status. Testing the interaction terms using the HT estimator is impossible since the number of endogenous variables increases with the inclusion of the interaction terms while there is no change in the number of available instruments. For men, the only interaction term that is significant is the Wave 2 time dummy variable[9]. This could indicate a true year effect or it could be a product of some data feature such as wage imputation for wave 2. The interaction term between wave 2 and immigrant status is included in all the panel regressions for men. We find no evidence that other variables affect immigrants and the native-born differently. McDonald and Worswick (1999) find the same. In the sample of women, returns to work experience and its square appear to vary between immigrants and the native-born. Interaction terms for experience and its square with immigrant status are thus included in all panel regressions for women. Interaction terms will need to be taken into account when interpreting the immigrant wage gap and assimilation effects.

Deciding which of the included variables are endogenous is of particular importance, as specifying the wrong instruments will lead to inconsistent and biased results for the HT estimator. Baltagi et al. (2003) provide a testing procedure, based upon the Hausman (1978) test, to determine the suitability of the HT estimator. They suggest a first Hausman test to distinguish between the random effects model and the fixed effects model. If the random effects model is rejected then a second Hausman test is carried out contrasting the HT estimator and the fixed effects model. The fixed effects model provides a suitable benchmark to test the exogeneity assumptions of the HT estimator. Failure to reject that fixed effects and HT estimators are identical can be interpreted as evidence that the assumptions of the HT estimator are valid. Hence, the choice of endogenous variables for the HT estimator can be tested using a Hausman test for the HT estimator versus fixed effects. We use experience, experience squared, the education dummies and immigrant status as potentially correlated with the individual effects. Intuitively, when we think of unobservable individual effects, we often think of ability and motivation both of which would affect the education level of an individual. More motivated individuals are likely to have greater work experience. Immigrants are likely to differ from native-born individuals in both ability and motivation; it is also possible that immigrants arriving at different points in time differ from one another in unobservable characteristics.

4.3 Instruments for time-invariant endogenous variables

[9] Dropping wave 2 data or dropping wave 1 and wave 2 data has no effect on the reported results.

Weak instruments can cause problems for any instrumental variable method. Statistically insignificant estimates and large standard errors for the time-invariant endogenous variables are obtained when using the HT estimator with weak instruments. In Table 4, we present the F-stat for the regression of each of the included endogenous variables on the time-varying exogenous variables (see Table 3) used to construct the instruments.

[Table 4 about here]

Staiger and Stock (1997) suggest that an F-stat less than 10 is associated with weak instruments. The F-stats in Table 4 are all greater than 10, although the F-stat for 'Certificate' in the female sample is close to 10. From the F-stats and correlations the instruments used for the remaining endogenous variables appear adequate and the time-varying endogenous variables are mean differenced to remove any unobserved individual effects. It is important to note that the coefficient estimates for other included variables are not affected by any inconsistency in the estimates of the non time-varying education variables.

5 Results

We estimate four models:

1. Heckman selection model on pooled panel data
2. Fixed effects
3. Random effects
4. Hausman-Taylor (IV)

For each of these we estimate a version of the model where we combine all immigrants and a version where we allow different initial wage gaps and assimilation profiles for immigrants from ESB and NESB. The purpose of the Heckman (1979) selection model is to provide a benchmark for comparison with previous Australian studies and with our panel estimates. Although the sample selection term is significant, the initial wage gap and assimilation profiles that we estimate if we use a linear regression model are almost identical to those from the Heckman model.[10]

[10] We get similar assimilation profiles if we use the same sample as that used for the model presented in Table A2. We also get similar profiles if we use an expanded sample where we include those observations which were previously excluded due to missing partner information.

The estimated relationships between characteristics and wage are in line with other Australian estimates. We focus our discussion on the estimates of the immigrant wage gap and assimilation profile. Our key results are summarised in Table 5. Appendix Table A2 contains details of the Heckman selection model and the estimates for panel models with different wage gap and assimilation profile coefficients for ESB and NESB immigrants are provided in Appendix Table A3.

[Table 5 about here]

Figure 1 shows the initial wage gap and wage assimilation profiles from the three models when we pool all immigrants together. Overall, controlling for unobserved heterogeneity produces a larger wage gap but shorter time to reach wage parity compared to the pooled Heckman model. The confidence intervals around these profiles are fairly wide (we suppress them in the graph for clarity) so at many points they are not significantly different from one another.

From Table 5, we can see that controlling for unobserved heterogeneity affects the estimates for men more than it does for women. When we compare the HT estimator to the pooled Heckman estimator, we see that for all men the initial wage gap nearly doubles (from 10.5 per cent to 20.2 per cent) and for ESB immigrant men the wage gap goes from zero to almost 20 per cent and becomes statistically significant. For NESB immigrant men, the wage gap in the pooled Heckman model is 19.6 per cent which increases to 23 per cent when we control for unobserved heterogeneity, but this difference is not significant. For all groups of women the change in the initial wage gap when we go from the pooled Heckman to the HT estimator is much smaller and the gap actually goes down for NESB immigrant women. For ESB immigrant women we find statistically insignificant wage gaps of 1 per cent (pooled Heckman model) and 6.4 per cent (HT model). For NESB immigrant women, we find a statistically significant wage gap of 19.6 per cent (pooled Heckman model) which decreases to 14.7 per cent, but remains significant, in the HT model. The fact that accounting for unobserved heterogeneity affects men more than women could be because we use a sample of workers and labour force participation is smaller for women than for men. It is likely that women self-select into work on the basis of favourable characteristics, such as high levels of education or motivation. The additional controls for unobservables through the HT estimator for women do not seem to affect the results very much.

Clearly, NESB immigrants, men and women, face a statistically significant and similar wage disadvantage on arrival. Our results suggest that NESB immigrants experience slow wage assimilation as also found by Chiswick and Miller (1985) and Beggs and Chapman (1988). Our results for the pooled Heckman estimator are close to those of other Australian studies, e.g. McDonald and Worswick (1999).

The major difference that we find between the HT results and the pooled regression or random effects results is that we find a large and statistically significant initial wage gap for ESB immigrant men. This finding is consistent with positive correlation between the observable characteristics on which immigrants are selected and unobservable characteristics such as ability and motivation which might affect wages. ESB male immigrants reach wage parity with natives within 12 years according to these estimates. Figure 2 shows the wage gap and assimilation profiles for male immigrants from English speaking backgrounds. Only the HT estimate profile and initial wage gap are significantly different than zero.

5.1 Panel estimators

Table A3 presents the full estimation results from the panel regressions. We reject the random effects model compared to the fixed effects model. We fail to reject that the HT and fixed effects models are the same. As discussed above, in the absence of model mis-specification this result is generally interpreted as rejecting the assumption that the unobserved effects are uncorrelated with the included variables (meaning that the random effects model is mis-specified) but that the exogeneity assumptions of the HT estimator are valid. Of course, the Hausman test procedure is known to be sensitive to general model mis-specification and the results should be taken with some caution. Nonetheless, this provides at least some evidence that the HT estimator controls for unobservable individual effects in the wage equations and employs acceptable exogeneity assumptions.

The HT estimate for the coefficient of tertiary education is quite high and statistically significant but it is not precisely measured. Comparable Australian studies that have controlled for unobservable individual heterogeneity using panel data are not available to contrast the size of the tertiary variable, but García-Mainar and Montuenga-Gómez (2005) obtained similarly large estimates for returns to education in Spain and Portugal using the HT estimator in a different context.

Finally, it is worth noting that the variance of the HT estimator is much higher than that of the fixed or random effects model. The standard error on the initial wage gap, for example, more than doubles. This is typical in instrumental variable estimation. One implication is that the confidence intervals on the initial wage gap and years to wage parity are quite large.

5.2 Cohort Effects

We re-estimate our panel data models allowing for different effects for different cohorts of immigrants. Our main motivation in doing this is for comparison with previous Australian studies that have examined cohort effects. There is a sense in which we do not observe cohorts. The oldest (first) cohort is comprised only of the individuals with the longest immigration history. We do not observe members of this first cohort in the years just after migration. In order to estimate cohort effects, we thus have to exploit the parametric specification—see section 5.3 below where we discuss and test this restriction further.

We consider five cohorts[11] based upon Commonwealth of Australia (2001) and the proposed labels therein: 'post-war resettlement' (1945-1965); 'a new emphasis for immigration policy' (1966-1977); 'population development is an important backbone of immigration policy' (1978-1984); 'economics is a major focus of immigration policy' (1985-1995); and 'migration planning becomes more focused on labour market issues' (1995 – present). We will refer to these cohorts as Cohort 1 through Cohort 5. Immigrant policy in Australia changed dramatically in the late 70s as the white-only immigration policy was gradually relaxed and the point system was introduced. Recently, greater emphasis has been placed on skilled migrants, a more racially equitable policy and accepting immigrants from any country provided they meet certain skill or humanitarian criteria.

[11] In an earlier version of the paper we considered three cohorts following Doiron and Guttmann (2009). These corresponded to grouping the first and second cohorts and the third and fourth cohorts mentioned here. The overall results are the same—newer cohorts appear to be of higher quality.

As shown in Table 6, immigrants that arrived later are better educated and from more diverse backgrounds than earlier immigrants. The majority of immigrants in the first three cohorts are from ESB as expected given the white Australia policy, which was in force before the 1970s. Table 7 presents results for the wage equations with cohort dummy variables.[12] Note that the cohort dummies are intuitively the same as generating an estimate of the average fixed effect by immigrant arrival group. Since the HT estimator, like a standard fixed effects model, allows for any distribution of the unobserved effects, different mean effects across different immigrant cohorts will not bias our HT estimates discussed above.

[Tables 6 and 7 about here]

Sufficient instruments are not available to estimate an HT model which simultaneously includes cohort effects and splits ESB/NESB immigrants. Thus we group all immigrants together, keeping in mind the results from Table 5 which show the effect of grouping all immigrants and separating them by language background.

Unlike previous studies (McDonald and Worswick, 1999, Miller and Neo, 2003), we do find evidence that cohort effects are present for both men and women, with immigrants who arrive later having a much smaller wage gap upon entry compared to earlier cohorts. For men and women, when we use the HT estimator, the initial wage gap is significantly different than zero for the cohorts that arrived before 1985. Immigrant men in Cohorts 1 through 3 earn, on average, 81% less than similar native-born Australians. Immigrant women in the first three cohorts earn 79% less, on average, than similar native-born Australians. Successive cohorts are better off, with the fourth and fifth cohorts facing a much smaller wage gap at entry than earlier cohorts. This is not surprising given that Australia's immigration policy is now more geared towards skilled immigrants than it was in the 1970s. The Australian labour market has become more open since the 1980s, with a move from centralised wage setting to enterprise bargaining and a decline in unionization rates. Improved economic conditions since the 1980s would have affected the relative returns to migration and increased positive selectivity of immigrants.

[12] As above, returns to experience are allowed to vary for immigrant and non-immigrant women but we find that the interaction terms in the HT estimates are jointly insignificant. Dropping these interaction terms does not change the fundamental conclusions.

Differences between NESB and ESB immigrants are present across all five cohorts. Note that if the cohort effects were generated by differences in the composition of immigrants from English and non-English speaking backgrounds, we would expect *smaller* cohort effects for earlier cohorts (which had a larger proportion of ESB immigrants). This is the opposite of what we find. The cohort effects are clearly not driven by English speaking background.

Although we find cohort effects, it is not possible to separate out exogenous changes in cohort quality from policy-induced effects. Both explanations are plausible. Finally, we also test a model where we include cohort effects for native born workers akin to Green and Worswick (2012). We use five cohorts matching those used for immigrants but based upon year of entry into the labour market. The estimated coefficients are statistically insignificant for all native cohorts; we thus prefer the specification reported in Table 7.

5.3 Sensitivity to parametric specification

One drawback of our approach is that we observe almost no immigrants at the actual point of migration[13]. Estimates of the initial wage gap and assimilation profiles are constructed by estimating the quadratic relationship between years of migration and wages on a set of immigrants who migrated in the past. While the range of years since migration used in estimation is therefore restricted, the coefficient estimates are used to extrapolate the relationship between years since migration and wages backwards to the point of initial migration. This problem is exacerbated in the cohort estimates since we are further shortening the range of years since migration on which the assimilation profile is estimated for each cohort and we are assuming that the relationship between wages and the other variables is identical across cohorts.

In both cases, if this parametric specification and these parameter restrictions are correct, then this approach is justified. Restricting values of the independent variable in a regression leads to higher standard errors (relative to some ideal where the variable is distributed continuously along the real line) but does not lead to bias. Extending a regression line beyond the observable range of data is always

[13] Prior to the top-up sample which HILDA conducted in 2011 (and which is not included in our data) the only way a new immigrant would enter the data is by partnering with a HILDA sample member.

questionable because it requires the parametric specification to hold not only within the range of observable data but also beyond what can be observed.

We cannot solve this problem with our data, but we can explore alternative choices of parametric specification to see whether or not our results are sensitive to the choice of the quadratic functional form which we use above. We explore three alternative specifications by replacing the quadratic in years since migration with: (a) natural log of years since migration; (b) a quartic in years since migration; (c) and the level of years since migration and the multiplicative inverse of years since migration.

These alternative specifications give qualitatively similar results, although both the initial wage gap and years to assimilation vary across specification. For example, for ESB immigrant men, when using natural log of years since migration we get an initial gap of 30 per cent (instead of 20) and 22 years to assimilate instead of 12. While these point estimates are quite different, the confidence intervals for these two estimates have substantial overlap and both point estimates are included in both confidence intervals.

The key patterns which we observe in Table 5 are unchanged in these alternative parametric specifications: extremely long assimilation times for both male and female, non-English speaking migrants, no gap for female migrants from English speaking countries, and a small wage gap and significant time to wage parity for male, English speaking migrants.

6 Discussion and Conclusions

In this paper we have attempted to improve our understanding of the immigrant wage gap and immigrant wage assimilation in Australia by estimating models which use panel data to control for unobserved differences between migrants and non-migrants. Most of our results are consistent with the previous Australian literature.

We find two novel results. First, we find that once we control for unobserved effects, there appears to be a wage gap on entry for male immigrants from English speaking backgrounds compared to native-born Australians. Other studies have failed to find such a gap. This result is not surprising if unobserved characteristics are positively correlated with observed characteristics. Since Australia's immigrants are selected on observable characteristics such as education it is not surprising that there is positive selection on unobservables such as ability and motivation as well.

Our second novel result is the finding of cohort effects. In particular, we find that more recent cohorts of immigrants appear to have smaller wage gaps than those from previous cohorts. The progressively better labour market performance of immigrants that arrive in later cohorts may be due to changes in Australian immigration policy that favours skilled migrants. It may also be due to the increased selectivity of immigrants. Economic and labour market conditions in Australia may be affecting the potential returns to migration and making Australia a more lucrative destination than in the past. Finally, recent improvements in source country conditions, such as in China, might make immigrants more work-ready in Australia than previously. Controlling for unobserved heterogeneity appears to be important in identifying these cohort effects.

Our results appear to fit with recent international studies (see section 2 above) which find that accounting for unobservable characteristics provides a less optimistic view of immigrant wage assimilation. Our finding of a larger wage gap and longer assimilation time for male, ESB immigrants when accounting for unobservables is also found in Canada, Germany and the U.S. It is difficult to know what policy implications to draw without having studied the effect of interventions on immigrant assimilation. One conclusion is that the current Australian policy of selecting on observable characteristics is even more selective than it appears on paper due to the positive correlation between observable and unobservable characteristics.

Earlier, we pointed out two data limitations and we return to these now. Higher attrition and lower response rates for immigrants may cause our estimates of the wage gap to be positively biased and result in under-estimating the time it takes for immigrants to assimilate to native-born Australians if the immigrants who remain in the sample are positively selected. From Table A1 in the appendix, however, such effects would appear to be small. Looking at education levels, we do not find that the selectivity of attrition differs between immigrant men from English speaking backgrounds as opposed to native Australians. This would not seem to be driving our finding of a wage gap for men from English speaking backgrounds. Our estimate of the wage gap for immigrants from non-English speaking backgrounds is more likely to be slightly overstated.

The fact that we are pooling immigrants who arrive as skilled migrants with those who are not vetted on selective immigration criteria will tend to bias our estimates towards zero, particularly if we are interested in the relationship between unobserved characteristics and observed characteristics for skilled migrants. Although we do not know the relative size of these two sources of bias (from attrition and from pooling heterogeneous visa types), they will tend to counteract each other. Without putting excessive emphasis on the actual point estimates which are presented, we are confident that our results are suggestive of a role for unobservable characteristics. As such, they should be taken into account when trying to understand wage assimilation patterns of immigrants to Australia.

References

Amemiya, T. and MaCurdy, T. E. (1986), 'Instrumental-variable estimation of an error-components model', *Econometrica*, 54(4), 869-880.

Antecol, H., Cobb-Clark, D. and Trejo, S. (2003), 'Immigration policy and the skills of immigrants to Australia, Canada, and the United States', *Journal of Human Resources*, 38(1), 192-218.

Antecol, H., Kuhn, P. and Trejo, S. (2006), 'Assimilation via prices or quantities? Sources of immigrant earnings growth in Australia, Canada, and the United States', *Journal of Human Resources*, 41(4), 821-40.

Australian Bureau of Statistics (ABS) (2009), 'A picture of the nation: the Statistician's report on the 2006 census', Commonwealth of Australia, Cat. no. 2070.0.

Australian Bureau of Statistics (ABS) (2011), 'Migration, Australia 2009-2010', Commonwealth of Australia, Cat no. 3412.0.

Baltagi, B. H. (2005), *Econometric analysis of panel data.* Wiley and Sons, Hoboken, NJ.

Baltagi, B. H., Bresson, G. and Pirotte, A. (2003), 'Fixed effects, random effects or Hausman-Taylor?: A pretest estimator', *Economics Letters,* 79(3), 361-369.

Baltagi, B. H. and Khanti-Akom, S. (1990), 'On efficient estimation with panel data: An empirical comparison of instrumental variables estimators', *Journal of Applied Econometrics,* 5(4), 401-406.

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

Beggs, J. J. and Chapman, B. (1991), 'Male immigrant wage and unemployment experience in Australia' NBER Chapters in: *Immigration, Trade and the Labour market*, 369-384, National Bureau of Economic Research Inc.

Beggs, J. J. and Chapman, B. J. (1988), 'Immigrant wage adjustment in Australia: Cross section and time series estimates', *Economic Record,* 64(3)**,** 161-167.

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. (1989), 'Immigrant and emigrant earnings: a longitudinal study', *Economic Inquiry,* 27(1)**,** 21-37.

Breusch, T. S., Mizon, G. E. and Schmidt, P. (1989), 'Efficient estimation using panel data', *Econometrica***,** 57(3), 695-700.

Chapman, B. J. and Iredale, R. R. (1993), 'Immigrant qualifications: Recognition and relative wage outcomes', *International Migration Review***,** 27(2), 359-387.

Chapman, B. J. and Mulvey, C. (1986), 'An analysis of the origins of sex differences in Australian wages', *Journal of Industrial Relations,* 28(4)**,** 504-520.

Chiswick, B. and Miller, P. (1995), 'The endogeneity between language and earnings: International analysis', *Journal of Labor Economics,* 13(2)**,** 246-288.

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

Chiswick, B. R., Lee, Y. L. and Miller, P. W. (2005), 'Immigrant earnings: a longitudinal analysis', *Review of Income and Wealth,* 51(4)**,** 485-503.

Chiswick, B. R. and Miller, P. W. (1985), 'Immigrant generation and income in Australia', *Economic Record,* 61(173)**,** 540-553.

Chiswick, B. R. and Miller, P. W. (2010), 'The effects of educational occupational mismatch on immigrant earnings in Australia, with international comparisons', *International Migration Review,* 44(4)**,** 869-898.

Cobb Clark, D. A. (2001), 'The longitudinal survey of immigrants to Australia', *Australian Economic Review,* 34(4)**,** 467-477.

Cobb-Clark, D. A. (2003), 'Public policy and the labor market adjustment of new immigrants to Australia', *Journal of Population Economics,* 16(4)**,** 655-681.

Cobb-Clark, D. A. and Connolly, M. D. (1997), 'The worldwide market for skilled migrants: can Australia compete?', *International Migration Review***,** 31(3), 670-693.

Cobb-Clark, D. A., Hanel, B. and Mcvicar, D. (2012), 'Immigrant wage and employment assimilation: A comparison of methods', IZA discussion paper No. 7062.

Commonwealth of Australia (2001), 'Immigration: Federation to Century's End 1901-2000', Statistics Section, Department of Immigration and Multicultural Affairs, Australia. Available at http://www.immi.gov.au/media/publications/statistics/federation/federation.pdf. Last viewed 15 July, 2013.

Commonwealth of Australia (2013), 'Australia's Migration Trends 2011–12', Department of Immigration and Citizenship. Economic Analysis Unit, Strategic Policy and Innovation Division, Australia. Available at http://www.immi.gov.au/media/publications/statistics/immigration-update/australian-migration-trends-2011-12.pdf. Last viewed 15 July, 2013.

Cornwell, C. and Rupert, P. (1988), 'Efficient estimation with panel data: An empirical comparison of instrumental variables estimators', *Journal of Applied Econometrics,* 3(2)**,** 149-155.

Doiron, D. and Guttmann, R. (2009), 'Wealth distributions of migrant and Australian born households', *Economic Record,* 85(268)**,** 32-45.

Fertig, M. and Schurer, S. (2007), 'Labour market outcomes of immigrants in Germany: The importance of heterogeneity and attrition bias', IZA discussion paper No. 2915.

García-Mainar, I. and Montuenga-Gómez, V. M. (2005), 'Education returns of wage earners and self-employed workers: Portugal vs. Spain', *Economics of Education Review,* 24(2)**,** 161-170.

Green, D. A. and Worswick, C. (2012), 'Immigrant earnings profiles in the presence of human capital investment: measuring cohort and macro effects', *Labour Economics***,** 19(2), 241-259.

Hausman, J. A. (1978), 'Specification tests in econometrics', *Econometrica***,** 46(6), 1251-1271.

Hausman, J. A. and Taylor, W. E. (1981), 'Panel data and unobservable individual effects', *Econometrica***,** 49(6), 1377-1398.

Heckman, J. (1979), 'Sample selection bias as a specification error', *Econometrica***,** 47(1), 153-161.

Hu, W. Y. (2000), 'Immigrant earnings assimilation: Estimates from longitudinal data', *American Economic Review,* 90(2)**,** 368-372.

Hum, D. and Simpson, W. (2004), 'Reinterpreting the performance of immigrant wages from panel data', *Empirical Economics,* 29(1)**,** 129-147.

Langford, M. S. (1995), 'The gender wage gap in the 1990s', *Australian Economic Papers,* 34(64)**,** 62-85.

Lubotsky, D. (2007), 'Chutes or ladders? A longitudinal analysis of immigrant earnings', *Journal of Political Economy,* 115(5)**,** 820-867.

Mcdonald, J. T. And Worswick, C. (1999), 'The earnings of immigrant men in Australia: assimilation, cohort effects, and macroeconomic conditions', *Economic Record,* 75(228)**,** 49-62.

Melbourne Institute of Applied Economic And Social Research (2010), 'HILDA Annual Report 2010', Available at http://melbourneinstitute.com/downloads/hilda/areport2010.pdf. Last viewed 3 November 2011.

Messinis, G. (2009), 'Earnings and languages in the family: Second generation Australians', *Economic Record,* 85(S1)**,** S59-S73.

Miller, P. and Neo, L. M. (2003), 'Labour market flexibility and immigrant adjustment', *Economic Record,* 79(246)**,** 336-356.

Preston, A. (2001), *The structure and determinants of wage relativities: evidence from Australia.* Ashgate, Burlington, VT.

Richardson, S. and Lester, L. (2004), 'A comparison of Australian and Canadian immigration policies and labour market outcomes', The National Institute of Labour Studies, Flinders University. Available at: http://www.immi.gov.au/media/publications/pdf/comparison_immigration_policies.pdf.

Staiger, D. and Stock, J. (1997), 'Instrumental variables regression with weak instruments', *Econometrica,* 65(3)**,** 557-586.

Wilkins, R. (2003), 'Immigrant earnings adjustment: the impact of age at migration', *Australian Economic Papers,* 42(3)**,** 292-315.

Wooden, M. and Watson, N. (2007), 'The HILDA survey and its contribution to economic and social Research (So Far)', *Economic Record,* 83(261)**,** 208-231.

Table 1: Sample Size by Wave

		Wave								
		1	2	3	4	5	6	7	8	9
	Immigrants	677	593	564	524	521	520	480	448	443
Men	Native-Born	1,951	1,950	1,867	1,848	1,873	1,915	1,871	1,899	1,940
	n	2,628	2,543	2,431	2,372	2,394	2,435	2,351	2,347	2,383
	Immigrants	624	557	534	505	516	523	505	477	466
Women	Native-Born	1,960	1,877	1,880	1,829	1,894	1,932	1,928	1,918	1,891
	n	2,584	2,434	2,414	2,334	2,410	2,455	2,433	2,395	2,357

Table 2: Key variables: Panel Sample Wave 5 Mean and Standard Deviations

	Men		Women	
Sub Group	**Native-Born**	**Immigrants**	**Native-Born**	**Immigrants**
Hourly Wage	25.82 (14.51)	26.88 (13.45)	22.03 (12.44)	23.33 (11.98)
Age	39.65 (9.53)	42.89 (9.61)	40.51 (9.47)	42.78 (9.14)
Experience	21.26 (10.27)	23.01 (10.46)	18.76 (9.22)	20.43 (9.73)
Partnered	0.74	0.80	0.71	0.76
Years since migration	N/A	23.04 (12.97)	N/A	24.47 (13.48)
Indigenous	0.02	N/A	0.02	N/A
Second Generation Migrant	0.27	N/A	0.24	N/A
English speaking background	N/A	0.51	N/A	0.47
Non-English speaking background	N/A	0.49	N/A	0.53
City	0.63	0.80	0.62	0.79
Inner regional	0.25	0.13	0.26	0.15
Outer regional/remote/ very remote	0.12	0.07	0.13	0.06
Bachelor's or higher (tertiary)	0.26	0.37	0.33	0.40
Certificate	0.42	0.34	0.28	0.24
Year12	0.12	0.13	0.13	0.18
Year 11 or less	0.20	0.16	0.25	0.17
Observations	1,873	521	1,894	516

Note: Standard deviations of continuous variables are shown in brackets.

Table 3: List of variables included in panel regressions

The dependent variable is the natural log of hourly wage	
	Variable
Time Varying Exogenous:	
	(Age/100)^2
	Partnered
	Years since migration / 100
	(Years since migration / 100)^2
	Four geographical location dummies are included: 1. City 2. Inner regional 3. Outer regional / remote / very remote
	Wave dummies are included for all nine waves
Time Varying Endogenous:	
	Experience/100
	(Experience/100)^2
Time Invariant Exogenous:	
	Indigenous
Time Invariant Endogenous:	
	Immigrant
	Tertiary
	Certificate
	Year 12

Table 4: F-Stats from the regression of each of the variables on the time-varying exogenous variables

	Men	**Women**
Variable	**F-Stat**	**F-Stat**
Immigrant	11,194.97	12,389.90
Tertiary	67.14	51.71
Certificate	17.46	10.45
Year 12	27.00	17.26

Table 5: Summary of regression results

	All immigrant men		ESB immigrant men		NESB immigrant men	
	Initial wage gap	Years to wage parity	Initial wage gap	Years to wage parity	Initial wage gap	Years to wage parity
Pooled Heckman	**-10.5%**	31 years	0.6%	0	**-19.6%**	36 years
Random Effects	**-13.5%**	18 years	-5.3%	7 years	**-19.4%**	26 years
Hausman-Taylor (IV)	**-20.2%**	21 years	**-19.7%**	12 years	**-23.0%**	35 years

	All immigrant women		ESB immigrant women		NESB immigrant women	
	Initial wage gap	Years to wage parity	Initial wage gap	Years to wage parity	Initial wage gap	Years to wage parity
Pooled Heckman	**-11.3%**	19 years	1.2%	0	**-19.6%**	22 years
Random Effects	**-3.2%**	4 years	13.9%	0	**-12.5%**	8 years
Hausman-Taylor (IV)	**-7.1%**	9 years	6.4%	0	**-14.7%**	10 years

Notes: Coefficients in **bold** are statistically significant at the 5 per cent level.

Table 6: Variable means and standard deviations by arrival cohorts of immigrants

	Immigrant Men					Immigrant Women				
	Cohort 1	Cohort 2	Cohort 3	Cohort 4	Cohort 5	Cohort 1	Cohort 2	Cohort 3	Cohort 4	Cohort 5
Hourly Wage	26.40	27.94	25.48	27.33	26.90	22.77	24.96	23.48	24.03	22.43
	(14.911)	(15.639)	(12.520)	(15.728)	(16.197)	(12.189)	(17.548)	(11.882)	(14.895)	(10.245)
Age	50.98	45.05	42.27	40.29	36.11	50.83	45.09	40.30	41.48	35.36
	(5.440)	(8.471)	(9.748)	(9.208)	(8.080)	(5.425)	(8.080)	(9.171)	(8.330)	(7.719)
Experience	32.51	26.43	22.71	19.46	15.29	27.78	22.70	18.39	18.92	13.00
	(6.904)	(9.563)	(9.893)	(9.588)	(8.482)	(7.702)	(9.015)	(8.921)	(8.909)	(8.027)
Years since migration	45.85	33.14	23.40	15.29	5.74	45.49	33.70	23.70	15.38	5.52
	(5.310)	(4.002)	(3.103)	(4.094)	(3.155)	(5.226)	(4.055)	(3.058)	(3.989)	(3.058)
English speaking background	0.58	0.59	0.53	0.41	0.46	0.54	0.60	0.48	0.35	0.43
Non-English speaking background	0.42	0.41	0.47	0.59	0.54	0.46	0.40	0.52	0.65	0.57
Bachelor's or higher	0.27	0.27	0.26	0.48	0.50	0.30	0.30	0.41	0.44	0.52
Certificate	0.49	0.36	0.39	0.29	0.25	0.25	0.29	0.23	0.25	0.27
Year 12	0.06	0.15	0.16	0.11	0.14	0.12	0.19	0.24	0.18	0.14
Year 11 or less	0.18	0.22	0.19	0.12	0.12	0.34	0.22	0.12	0.13	0.08
Observations	515	1,135	766	1,538	816	689	1,098	811	1,425	684

Notes:

(i) Standard deviations are in parenthesis. Standard deviations are not provided for dummy variables.

(ii) Definition of arrival cohorts are: cohort 1 arrived before 1966; cohort 2 arrived between 1966 and 1977; cohort 3 arrived between 1978 and 1984; cohort 4 arrived between 1985 and 1995; and cohort 5 arrived after 1995.

Table 7: Fixed effects, random effects and HT(IV) estimates of wage equations with immigrant cohort effects

	Men			Women		
VARIABLES	Fixed Effects	Random Effects	HT(IV)	Fixed Effects	Random Effects	HT(IV)
(Age/100)^2	-2.91**	-3.07***	-4.06***	-3.43***	-0.61***	-2.85***
	(1.313)	(0.799)	(0.801)	(0.879)	(0.108)	(0.535)
Experience/100	5.61***	2.59***	4.68***	3.68***	2.66***	4.11***
	(1.006)	(0.355)	(0.507)	(0.863)	(0.190)	(0.489)
(Experience/100)^2	-2.28*	-1.60**	-1.22	-1.15	-3.98***	-1.77
	(1.296)	(0.715)	(0.842)	(1.001)	(0.441)	(0.734)
Immi*(Experience/100)				-2.20	-1.34***	-1.85**
				(1.653)	(0.391)	(1.018)
Immi*(Experience/100)^2				2.78*	2.55***	2.93**
				(1.493)	(0.920)	(1.359)
Partnered	0.01	0.04***	0.01	0.01	0.03***	0.01
	(0.009)	(0.008)	(0.008)	(0.010)	(0.008)	(0.009)
Immi*Cohort01		-1.87	-3.01***		-0.57	-1.76*
		(1.145)	(1.120)		(1.030)	(1.070)
Immi*Cohort02		-2.17***	-2.35***		-0.91	-1.23**
		(0.571)	(0.564)		(0.601)	(0.625)
Immi*Cohort03		-1.88***	-1.85***		-1.89***	-1.79***
		(0.546)	(0.539)		(0.562)	(0.557)
Immi*Cohort04		-0.26**	-0.35		-0.16	-0.13
		(0.105)	(0.176)		(0.119)	(0.138)
Immi*Cohort05		-0.12***	-0.22		-0.08	-0.05
		(0.044)	(0.175)		(0.053)	(0.106)
C01*(Years since migration/100)	14.34***	8.61*	14.28***	9.06*	3.75	9.15*
	(5.484)	(5.023)	(4.873)	(5.312)	(4.546)	(4.790)
C01*(Years since migration/100)^2	-16.73***	-9.64*	-16.57***	-9.47	-4.52	-10.14*
	(6.034)	(5.482)	(5.358)	(5.766)	(4.981)	(5.225)
C02*(Years since migration/100)	14.24***	13.20***	14.28***	8.28*	5.44	7.14**
	(3.675)	(3.422)	(3.266)	(4.200)	(3.567)	(3.685)
C02*(Years since migration/100)^2	-21.05***	-19.49***	-21.06***	-9.25	-6.42	-8.90
	(5.445)	(5.097)	(4.841)	(5.792)	(5.256)	(5.301)
C03*(Years since migration/100)	16.18***	16.21***	16.22***	16.53***	16.85***	16.30***
	(4.946)	(4.657)	(4.397)	(5.256)	(4.766)	(4.699)
C03*(Years since migration/100)^2	-34.88***	-34.50***	-34.85***	-32.70***	-34.56***	-32.91***
	(10.419)	(9.827)	(9.263)	(10.636)	(10.029)	(9.743)
C04*(Years since migration/100)	2.64*	2.82**	2.71*	3.75*	2.92*	3.41**
	(1.472)	(1.353)	(1.307)	(2.109)	(1.511)	(1.653)
C04*(Years since migration/100)^2	-7.11	-7.44*	-7.12	-7.78	-7.34	-7.97*
	(4.681)	(4.345)	(4.161)	(5.173)	(4.813)	(4.738)
C05*(Years since migration/100)	0.74	0.39	0.74	5.37**	3.86**	5.09***
	(1.390)	(1.292)	(1.235)	(2.267)	(1.511)	(1.815)
C05*(Years since migration/100)^2	3.78	6.79	3.98	-30.12**	-24.80**	-30.35**
	(10.640)	(10.078)	(9.455)	(13.672)	(11.999)	(12.517)
Indigenous		0.00	0.11		0.01	0.08
		(0.046)	(0.231)		(0.039)	(0.082)
Tertiary		0.48***	1.28***		0.38***	0.45
		(0.018)	(0.433)		(0.014)	(0.195)
Certificate		0.14***	0.10		0.09***	-0.11
		(0.015)	(1.026)		(0.014)	(0.453)
Year 12		0.15***	0.34		0.11***	0.12
		(0.021)	(1.374)		(0.017)	(0.687)
Immi * wave2	-0.05***	-0.04***	-0.05***			
	(0.015)	(0.015)	(0.013)			
Constant	2.17***	2.31***	2.30***	2.63***	2.46***	2.55***
	(0.279)	(0.107)	(0.725)	(0.195)	(0.022)	(0.321)
Hausman Test		126.58	1.40		54.04	4.18
p-value		0.0000	1.0000		0.0003	1.0000

Notes: (i) Wave dummies and location variables were also included.
(ii) For men, interaction terms were included for wave 2 and immigrant status.
(iii) For women, interaction terms were included for experience and its square and immigrant status.
(iv) Standard errors are in parentheses.
(v)*** p<0.01, ** p<0.05, * p<0.1

Table A1:
Educational background of sample leavers and stayers by immigrant status

	Sample Stayers	Sample Leavers	Sample Stayers	Sample Leavers	Sample Stayers	Sample Leavers
			Men			
	Immigrants				Non-immigrants	
Education	ESB	ESB	NESB	NESB		
More than high school	1196 (71%)	333 (70%)	1204 (73%)	346 (66%)	8460 (69%)	1916 (65%)
High school or less	494 (29%)	141 (30%)	435 (27%)	178 (34%)	3755 (31%)	1043 (35%)
Sample size	1690	474	1639	524	12215	2959
			Women			
	Immigrants				Non-immigrants	
Education	ESB	ESB	NESB	NESB		
More than high school	1041 (67%)	280 (63%)	1085 (63%)	325 (61%)	7368 (61%)	1814 (58%)
High school or less	505 (33%)	167 (37%)	628 (37%)	210 (39%)	4726 (39%)	1310 (42%)
Sample size	1546	447	1713	535	12094	3124

Notes: Table is created by pooling across all waves. "Sample Stayers" are those present in wave *t* and wave *t+1;* "Sample Leavers" are those present in wave *t* but absent from survey in wave *t+1*.

Table A2: Key coefficient estimates from baseline model. Heckman sample selection model for log hourly wage on immigrant status

	Men		Women	
Variable	(1)	(2)	(1)	(2)
ESB		0.01		0.01
		(0.055)		(0.052)
ESB*(Years since migration/100)		0.07		0.08
		(0.468)		(0.428)
ESB*(Years since migration/100)^2		-0.20		-0.00
		(0.885)		(0.807)
NESB		-0.20***		-0.19***
		(0.055)		(0.046)
NESB*(Years since migration/100)		0.92*		1.22***
		(0.506)		(0.404)
NESB*(Years since migration/100)^2		-1.00		-1.57**
		(1.039)		(0.733)
Immigrant	-0.11***		-0.11***	
	(0.041)		(0.036)	
Years since migration/100	0.55		0.76***	
	(0.349)		(0.295)	
(Years since migration/100)^2	-0.66		-0.91*	
	(0.682)		(0.546)	
Lambda	-0.19***	-0.19***	0.02**	0.02***
	(0.036)	(0.038)	(0.012)	(0.012)
Sample size (individuals)	22,466 (5,332)		24,447 (5,551)	

Notes:
(i) The model includes quadratics in age and experience, three education dummies, partnered and indigenous indicators, location (urban, inner regional, outer regional, remote) dummies and time dummies. The selection equation includes those variables and whether the household has an outstanding mortgage, the partner's wage, and dummy variables for the presence in the household of children in the 0-4 age and 5-14 age ranges.
(ii) Maximum likelihood estimates and clustered standard errors.
(iii) Standard errors are in parentheses.
(iv) *** p<0.01, ** p<0.05, * p<0.1

Table A3: Fixed effects, random effects and HT(IV) estimates of wage equations with ESB and NESB immigrant dummy variables

Variable	Men			Women		
	Fixed Effects	Random Effects	HT(IV)	Fixed Effects	Random Effects	HT(IV)
(Age/100)^2	-2.97**	-1.59***	-4.21***	-3.45***	-0.60***	-2.90***
	(1.314)	(0.219)	(0.759)	(0.877)	(0.108)	(0.578)
Experience/100	5.62***	3.14***	4.68***	3.69***	2.65***	4.10***
	(1.006)	(0.184)	(0.437)	(0.863)	(0.190)	(0.467)
(Experience/100)^2	-2.31*	-2.78***	-1.19	-1.15	-4.00***	-1.75**
	(1.297)	(0.382)	(0.809)	(1.000)	(0.440)	(0.759)
Immi*(Experience/100)				-1.84	-1.45***	-1.03
				(1.651)	(0.389)	(0.787)
Immi*(Experience/100)^2				2.39	2.79***	2.46*
				(1.471)	(0.911)	(1.354)
Partnered	0.01	0.04***	0.01	0.01	0.03***	0.01
	(0.009)	(0.008)	(0.008)	(0.010)	(0.008)	(0.009)
ESB		-0.05	-0.20*		0.14***	0.05
		(0.044)	(0.106)		(0.053)	(0.117)
ESB*(Years since	2.11***	0.89**	2.13***	0.70	0.00	-0.05
migration /100)	(0.568)	(0.343)	(0.506)	(1.495)	(0.359)	(0.734)
ESB*(Years since	-4.40***	-1.80***	-4.23***	1.08	0.38	0.59
migration /100)^2	(1.017)	(0.650)	(0.895)	(1.162)	(0.685)	(1.056)
NESB		-0.19***	-0.22**		-0.12***	-0.16
		(0.041)	(0.115)		(0.044)	(0.104)
NESB*(Years since	0.86	1.05***	0.91*	2.49*	1.71***	1.76**
migration /100)	(0.543)	(0.347)	(0.482)	(1.454)	(0.324)	(0.692)
NESB*(Years since	-0.86	-1.13*	-0.71	-3.14***	-2.21***	-2.82***
migration /100)^2	(1.078)	(0.683)	(0.943)	(1.112)	(0.597)	(0.962)
Indigenous		0.00	0.13		0.01	0.08
		(0.046)	(0.179)		(0.039)	(0.077)
Tertiary		0.48***	1.32***		0.38***	0.43**
		(0.018)	(0.344)		(0.014)	(0.219)
Certificate		0.14***	0.25		0.09***	-0.10
		(0.015)	(0.681)		(0.014)	(0.467)
Year 12		0.14***	0.11		0.10***	0.31
		(0.021)	(0.978)		(0.017)	(0.826)
Immi*wave2	-0.05***	-0.04***	-0.05***			
	(0.015)	(0.014)	(0.013)			
Constant	2.42***	2.50***	2.27***	2.81***	2.45***	2.53***
	(0.275)	(0.023)	(0.486)	(0.188)	(0.022)	(0.358)
Hausman Test		144.93	1.30		56.36	7.33
(*p*-value)		(0.000)	(1.000)		(0.000)	(0.996)
Sample size		21,884			21,816	
(individuals)		(4,928)			(4,894)	

Notes: (i) Wave dummies and location variables were also included. (ii) For men, interaction terms were included for wave 2 and immigrant status. (iii) For women, interaction terms were included for experience and its square and immigrant status. (iv) Standard errors are in parentheses. (v)*** p<0.01, ** p<0.05, * p<0.1