Skills and rural-urban wage differences in Australia

by

Phuong Ho

A Thesis in the Field of Economics

for the Degree of Master of Research in Economics

Macquarie University

27 March 2019

Table of Contents

Summary

Growing research on the urban wage premium shows that workers in urban areas earn more than workers with similar skill levels in rural areas. In Australia, little is known about whether the urban wage premium exists or the magnitude of the premium. Using a panel approach, the study finds that differences in cognitive ability and personality traits have little impact on rural-urban wage differentials. When other differences in individual characteristics are considered, Australian workers in large urban centers still earn around 7.5% more than workers in rural areas. The relationship between local economy size and local wages is robust when endogeneity issues are accounted for by instruments. It is not evidenced from the study that stayers in urban areas enjoy higher wage growth than stayers in rural areas as the learning hypothesis suggests. It is more likely that rural-to-urban migrants go through a period of social acclimatization when they do not receive a full urban wage premium upon arrival but experience high wage growth the following year. The analysis undertaken in this dissertation suggest that in Australia, like in other countries, how much we earn depends not only on our abilities but also external factors.

Statement of originality

This work has not previously been submitted for a degree or diploma in any university. To the best of my knowledge and belief, the thesis contains no material previously published or written by another person except where due reference is made in the thesis itself.

Phuong Ho 27 March 2019

Acknowledgements

To my grandfather Hanh Nguyen: you are my inspiration. You are the reason that I attend the program. Life is a journey, and I try to spend each minute to learn new things and to experience this wonderful adventure.

Special thanks to my mother, Nguyen Nguyen, and father, Tich Ho, for your unconditional beliefs in me and your encouragement over the years.

I am also grateful to invaluable advices throughout the process from Kompal Sinha and Craig MacMillan, my wonderful supervisors, and helpful comments from Frank Song and Jorge De la Roca. Karin Hosking provided professional copyediting assistance. I would like to give special thanks to Tony Bryant for the weekly meetings and his support for research students.

Lastly a big thank you to all my friends in the Economics class. I will always remember the times we spent together for the last two years. I wish all of you the best!

Without you it would not be possible!

1 Introduction and purpose

'Urbanization is not about simply increasing the number of urban residents or expanding the area of cities. More importantly, it's about a complete change from rural to urban style in terms of industry structure, employment, living environment and social security.'

Li Keqiang, Premier of the State Council of the People's Republic of China

More than half of the world's population is now living in urban areas. The urban population increased from less than 30% of the world's population in 1950 to 54% in 2014. About half of the world's urban residents live in small settlements with more than 500,000 people, and one in eight live in large cities with more than 10 million people (United Nations, 2014). While developed countries in North America and Europe have higher percentages of people living in cities, at around 70%, developing countries in Africa and Asia are catching up with higher rates of urbanization. The United Nations (2014) forecasted that by 2050, 66% of the world population would live in urban areas. Australia is one of the most urbanized countries in the word. As of 2001, eight in 10 Australians lived within 50 kilometers of the coastline, and most of them lived in capital cities situated near the coast (Australian Bureau of Statistics (ABS), 2004).

As a large portion of the population is concentrated in cities, cities are the hubs of economic and cultural activities. Sydney and Melbourne, the two largest cities in Australia, contributed 24.1% and 18.3% of the nation's GDP in the period 2015 to 2016, and on average 22.5% and 18% of the GDP growth in the period 1989 to 2016 (SGS Economics and Planning, 2015). The role of cities has been well recognized since ancient times. Tucker (1843, p. 127) wrote that the growth of cities marks 'the progress of intelligence and the arts, measures the sum of social enjoyment, and always implies excessive mental activity', and cities are the results to which all countries 'inevitably tend'.

Given the importance of cities, it is important to understand why people are attracted to cities and what the benefits are of gathering in a small area. One important aspect of urbanization that has been observed around the world is the urban wage premium, the phenomenon that workers in urban areas earn more than workers in other areas. In the US, urban workers earned 33% more than their rural counterparts in 1992 (data from Statistical Abstracts of the US, Glaeser & Maré, 2001). In France, for the period from 1996 to 1998, the average wages of workers in Paris were 15% higher than in other large French cities, 35% higher than in mid-sized cities and 60% higher than in rural employment areas (data from Annual Social Data Declarations, Combes, Duranton, & Gobillon, 2008). In Spain, workers in Madrid, the biggest city, earned 21% more than workers in Valencia, the third biggest city, and 55% more than workers in rural areas (De la Roca & Puga, 2017). The same pattern is also observed in Australia; workers in Australian capital cities have earned 17% to 22% more than workers in other parts of the states since 1995 (Australian Government — Treasury, 2017, based on ABS data).

Empirical studies on the topic, like that of Glaeser and Maré (2001), often aim to measure how much more an urban worker earns than a rural worker who has similar skill levels. Thus, the main interest lies in accounting for individual heterogeneity to isolate location effects on individual wages rather than simply comparing average wages between rural and urban areas. Answering whether spatial wage differentials are due to individual differences or differences in locations helps in directing research efforts. If the urban wage premium is due to skills or ability bias, research should focus on why more able workers are attracted to urban areas. If the premium is due to locations, research should focus on how urban areas affect local wages (Glaeser & Maré, 2001).

The existence of the urban wage premium suggests a positive relationship between individual incomes and local economy size: workers in large cities earn more than workers in medium-sized cities, and workers in medium-sized cities earn more than workers in rural areas, etc. The concept that growth in local economy size increases local workers' and firms' income is referred to as agglomeration economies (Combes & Gobillon, 2015). Agglomeration forces drive up wages in urban areas, areas with large economy sizes, creating the urban wage premium. Marshall (1890, p. 156) suggested that a concentration of workers with the same type of skills encourages learning: ideas are easy to spread, and efforts are more likely noticed and rewarded if there are more people nearby. Because of the acceleration of human capital in urban areas, workers in these areas have higher skill levels, and consequently they earn more than workers in other areas. Glaeser (1999), formularizing Marshall's explanation, proposed that positive association between density and learning would lead to high mean and high variance of skills and a high proportion of young people in cities. The prediction of high mean and high variance of skills in urban areas is in line with the empirical findings of Combes, Duranton, Gobillon and Roux (2012) for French cities. Together, the theories of Glaeser (1999) and Marshall (1890) and the empirical findings of Combes et al. (2012) suggest that densely populated areas result in high skills and high wages of workers in these areas, and that explains the positive relationship between local wages and the size of the local economy.

In Australia, there are few studies on the urban wage premium. Some related topics have included determinants of city growth in Australia (Bradley & Gans, 1998), job mobility in Australian metropolitan and non-metropolitan areas (Bill, Mitchell, & Welters, 2006), and the retention rates of Australian university graduates in major cities and regional areas (Corcoran, Faggian, & McCann, 2010). Studies on wage differentials in Australia, on the other hand, have focused mainly on the gender wage gap; examples include with the works of Kidd and Shanon (1996), Miller and Rummery (1991), Haig (1982) and Jones (1983).

One study on the topic in the country is of Rowe, Corcoran and Bell (2017). Using the 2003 Longitudinal Survey of Australian Youth (LSAY), the authors examined differentials in the entry-level wages of Australian non-metropolitan young migrants and non-metropolitan stayers. Non-metropolitan migrants on average earned 25% more than non-metropolitan stayers. The authors argued that initial wage loss and subsequent year wage gain following metropolitan migration implied that non-metropolitan migrants took advantage of opportunities in metropolitan areas that were not available in non-metropolitan areas.

It is not known whether workers in urban Australia earn a wage premium relative to workers with similar skills in the country's rural areas or how much the premium is. Given the lack of research interest in the topic, this study aims to provide better understanding of the urban wage premium in Australia. Like other empirical wage premium studies, the study primarily focuses on measuring the extent to which urban-rural status affects individual wages. This involves separating location and individual effects on individual wages. To that end, the study incorporates econometric developments in the field in accounting for individual heterogeneity. Further analysis explores whether agglomeration economies exist in Australia via examining the link between individual wages and local employment density. In contrast to Rowe et al. (2017), who compared wage outcomes of non-metropolitan stayers and non-metropolitan migrants and evaluated the effects of school-to-work pathways on young people's wages, the study focuses on accounting for individual differences to obtain reliable estimates of the urban wage premium.

In the larger context, urban research is necessary for Australia for several reasons. Australia ranks 6th in the world in area with more than 7.7 million sq. km and ranks 56th in population with 23 million people (CIA, 2016). Population growth remains strong, with 1.6% average annual growth for the 5-year period from 2011 to 2016, and the population is projected to double to 46 million by 2075 (ABS, 2013). Meanwhile, housing affordability has been decreasing since the 1980s due to 78% increase in housing prices in the period from 1980 to 2015 (OECD, 2018). High urbanization and future expansions highlight the importance of understanding individuals' location choices and density impacts on the local economy in Australia. Further research on urban topics such as the urban wage premium and agglomerations will help to answer important questions such as why do people favor one place over others? which forces drive productivity in the area? or which areas possess ideal characteristics for economic development? It also seems inadequate to study housing prices, a topic receiving a great deal of public and research attention, without understanding which factors attract people to an area, and how an area's population affects local wages, prices and housing costs.

2 Research contents

Using data from the HILDA survey, the Australian Bureau of Statistics (ABS) and other sources, the study

1. Examines how wage premiums are related to large urban centers and small urban centers in Australia. Individual heterogeneity is controlled for by the HILDA's individual characteristic variables. Incorporating different perspectives on skills, the study includes Mincerian human capital and occupational skill controls. The study also examines how individual cognitive ability and personality traits affect estimates of the urban wage premium. Unobserved individual differences are addressed using individual fixed effects, following Glaeser and Maré (2001).
2. Examines in more details how migration decisions, moving to other areas or staying at the same place, affect individual wages in the short run and long run (over several years). The analysis explores urban learnings, transitional noise and their effects on urban wage premium estimates. The analysis follows First Difference models and the 'long' difference models of Yankow (2006).
3. Answers whether agglomeration economies exist in Australia by applying Combes et al. (2008)'s two-step procedure.

This study focuses on full-time male workers and the period from 2001 to 2016. The following literature review is very far from a complete review on the urban wage premium, a broad and interesting topic, but provides the basis for the methods used in this study.

In addition to providing first empirical results of the urban wage premium in the case of Australia, the studies contribute to the literature on several other ways. The effects of personality traits on the urban wage premium are suggested (Combes & Gobillon, 2015), but have not been tested. The processes underlying high wage growth post rural-to-urban migration and why urban-to-rural migrants do not experience significant wage loss are not clear in the previous studies (Glaeser & Maré, 2001; Rowe et al., 2017; Yankow, 2006). These are the issues that the study aims to shed light on.

The first and second analyses divide Australia into large urban centers, small urban centers and rural areas based on the ABS's definition of 'urban'. Studying 'overall' urban effects is not a disadvantage because we would like to understand not only the effects of local economy size but also other urban characteristics such as the effect of airports, ports, hospitals, etc. on local wages. The third analysis divides Australia into labor market regions, focusing solely on the relationship between local economy size and local wages. Using two structures allows examination of urban wage premiums from different perspectives.

3 Literature review

This section provides the theoretical and empirical background for the study. The theory review highlights the important role of location in individual wages. For an empirical urban wage premium study like this one, however, the main concern is to disentangle individual effects and location effects on individual wages. Recent empirical research on the urban wage premium focuses on the use of panel data on individuals to control for individual differences, especially via the use of individual fixed effects (Glaeser & Maré, 2001). I mainly review econometric approaches using panel data for their above advantage and because they are directly relevant to this study, which uses the Household, Income and Labour Dynamics in Australia (HILDA) panel data.

3.1 Spatial equilibrium theories

In a spatial equilibrium, both individuals and firms have no incentives to relocate: individuals maximize their utility and firms maximize their profits by staying where they are. On the labor supply side, if there are no relocation barriers, identical workers in different locations should achieve the same utility level. If that is not the case, they will have incentives to move to areas where they attain higher utility. Potentially, there are many factors contributing to an individual's utility function such as his income, the surrounding environment and local foods. In a simple case where utility levels depend solely on the amount of goods and services consumed, equal utility for similar workers across locations is the same as equal real wages. In that case, two persons have different real wages when their individual characteristics are different. Differences in nominal wages among individuals result from differences in their real wages, or their characteristics, and differences in local prices. As a result, it is necessary to consider both individual characteristics and location characteristics when examining individual wages.

On the labor demand side, for a firm to stay in a high wage area, local workers should be more skilled, or there would be local factors that enhance the firm's production relative to other areas or allow the firm to charge higher prices for its products and services. In particular, assuming a Cobb—Douglas production function and a firm's zero economic profit, Glaeser and Maré (2001) shows that differences in workforce skills and location productivity explain differences in wages between areas. 'Location productivity' here consists not only of the 'real' total factor of productivity but also local nominal price levels. With similar assumptions and in a more detailed set-up, Combes et al. (2008) and Combes and Gobillon (2015) separated individual i's wages w_{it} in year t into 'composite' local productivity B_{ct} and individual skill s_{it}, $w_{it} = \left(\frac{A_{ct}p_{ct}}{{r_{ct}}^{1-\alpha}}\right)^{1/\alpha} s_{it} \equiv B_{ct}s_{it}$ or in log form $lnw_{it} = lnB_{ct} + lns_{it}$. Individual wages are divided into two parts, location part and individual part. Local composite productivity B_{ct} consists of local total factor of productivity A_{ct} and other 'price' factors, namely price of outputs p_{ct} and price of inputs r_{ct}. Like Glaeser and Maré (2001), in Combes et al. (2008)'s and Combes and Gobillon (2015)'s specifications, differences in composite productivities among areas could be purely nominal, reflecting differences in local price levels, rather than indicating differences in local productivity.

There is voluminous theoretical and empirical literature on how growth in local economy size affect the composite (local) productivity factor B_{ct}. The effects of the size on B_{ct} could go through either A_{ct}, p_{ct} or r_{ct}. For example, having many people means that cities can share the costs and enjoy the benefits of expensive facilities such as airports and hospitals (see Berglas and Pines (1981) for the relation between local public goods and the size of the sharing group and Abdel-Rahman and Fujita (1990) for sharing of a base of suppliers in cities). In that case, the costs of input r_{ct} decreases, increasing B_{ct}. As mentioned above, densely populated areas encourage individual learning (Glaeser, 1999; Marshall, 1890). As a result, workers in the area are more productive than workers in other areas, leading to high local A_{ct} and so high B_{ct}. Besides learning and learning hypotheses, 'matching' theories propose that density improves both the quantity and quality of matching among people and firms. For example, Coles and Smith (1998)'s setup of flows and stocks of buyers and sellers resulted in a probability matching function that has increasing returns to scale. In other words, employees are more likely to find a suitable job and employers are more likely to find workers with right skills if there are more matching opportunities in the areas.[1] In addition to agglomeration forces, big cities generate dispersion forces, the disadvantages that reduce B_{ct}, such as high crime and high stress levels people suffer from living in cities. If there are only advantages, nothing will stop cities increasing in size. If there are only disadvantages, cities will not exist. The size of a city strikes a balance between agglomeration and dispersion forces. As these forces are relevant in explaining B_{ct}, they are also relevant in explaining spatial wage differentials.

More importantly, the above arguments from both the labor supply side and the labor demand side suggest that both location characteristics and individual characteristics determine a worker's wages. This study's main goal is to estimate the overall location effects, lnB_c (assuming location effects are fixed over time), and their relation to local economy size rather than to identify the underlying processes. For clarity, an area's 'wage premium' that we will estimate is lnB_c, the percentage that workers in that area earn

[1] These are only a few examples of agglomeration forces. Duranton and Puga (2004) and Puga (2010) summarize research on the causes of agglomeration economies into three very broad categories of sharing, matching and learning.

more than workers with similar skill levels in the comparison area. lnB_c is also called location effects in this study. If there are only two broad categories of urban and rural areas, lnB_{Urban} is the urban wage premium.

3.2 Empirical methods

As individual wages can be divided into an individual part and a location part (Combes et al., 2008; Combes & Gobillon, 2015; Glaeser & Maré, 2001), individual heterogeneity needs to be appropriately accounted for to obtain accurate estimates of the urban wage premium. Among individual characteristics, determinants of individual skills are the most important as they decide how productive a worker is and so how much he earns.

3.2.1 Perspectives on skills

Skills in urban wage premium studies can be viewed as human capital levels in human capital theories. Combes et al. (2008, p. 726) termed skills as 'fixed individual attributes which are rewarded on the labor market'. Glaeser and Maré (2001) defined the skills of worker i as his efficiency unit labor $\emptyset_i$. His real wages would be $\emptyset_i w_c/p_c$ where w_c and p_c are local wages per efficiency unit and the price level at location c where worker i lives. With the utility equalization assumption, the real wages of workers with similar skills are equal over space, or $\frac{w_c}{p_c}$ is equal for all c. In that formulation, differences in real wages are solely resulted from differences in skills. A worker with two times the amount of 'skills' would earn two times the real wage. Human capital theories have a similar description of human capital K with earnings $E = wK$ where w is the rental rate of a unit of human capital (Mincer, 1974). The Mincer earnings function, based on human capital approaches, describes an individual's lifecycle earnings via education and work experience. In the Mincer earnings function, the log of wages equals the sum of years of schooling, years of work experience and years of work experience squared. Following that tradition, studies on spatial wage differentials often use individual wages in log form, and include some types of education and work experience variables (Di Addario & Patacchini, 2008; Duranton & Monastiriotis, 2002; Glaeser & Maré, 2001; Yankow, 2006). Considering nonlinear return to years of schooling, some studies include education levels (De la Roca & Puga, 2017; Di Addario & Patacchini, 2008).

As human capital accumulation is different for different demographic groups, it is reasonable to include these individual demographic variables in wage equations. Many studies are restricted to males (De la Roca & Puga, 2017; Glaeser & Maré, 2001; Glaeser & Resseger, 2010; Yankow, 2006). There are several reasons for the exclusion of females. First, years of work experience, derived via age and years of schooling, will overestimate females' work experience due to females' lower workforce participation.

Second, education and work experience may not have the same influences on males' and females' capital accumulation. Third, females may earn different wages to similar skilled males due to gender wage discrimination. De la Roca and Puga (2017) found that the city size premium for females was much lower than that for males: the medium-term earnings elasticity with respect to city size for female workers was 2.3% compared to 5.1% for male workers in Spain.

In addition to the traditional Mincerian controls of education and work experience, other studies include direct measures of individual skills such as academic grades and ability test scores. The argument is that education levels may not be the same as intelligence and capability and so not totally predict how well people do their jobs. Researchers using NLSY data often include the Armed Forces Qualification Test (AFQT) score as a direct measure of an individual's cognitive ability (Glaeser & Maré, 2001; Yankow, 2006; Gould, 2007). The AFQT consists of tests on 'arithmetic reasoning, mathematics knowledge, paragraph comprehension, and word knowledge', primarily for determining the US enlistment eligibility (ASVB, n.d.). Nevertheless, including the AFQT makes little differences to the estimates of the urban wage premium; in Glaeser and Maré (2001)'s and Yankow (2006)'s studies, the premium decreased by less than 1% after the AFQT variable is added.

Bacolod, Blum and Strange (2009)'s approach on skills is particularly interesting. They argued that a hedonic market clearing process matches a worker' skills to the occupation in which he is employed. As a worker possesses skills that enable him to do his job, his skills can be determined indirectly via his occupation. The three main types of skills, according to Bacolod et al. (2009), are cognitive skills, people skills and motor skills. Bacolod et al. (2009) used information from the Dictionary of Occupational Titles (DOT) to determine the level of each type of skills required for an occupation. For instance, physicists and life scientists were attributed with high cognitive skills while dentists and machinists were attributed with high motor skills. These DOT skills were then added to regression equations besides traditional education and work experience variables.

3.2.2 Individual fixed effects

Even with all the 'observed' skill differences from datasets, there is no guarantee that all individual differences are appropriately controlled for. Learning facilities in large cities potentially draw individuals with learning aptitude to the area. These skills and learning habits can be then transferred to their children, who likely live in the same city (Combes & Gobillon, 2015). Competitive, ambitious or extroverted people may prefer cities because of the broad range of career options, nice restaurants and busy nightlife in the area. These individual characteristics, at the same time, affect how much an individual earns.

Observed individual characteristics such as gender, age and education often available in datasets do not reflect all the differences among individuals, especially characteristics that are hard to quantify and measure like ambition or personality. Omitting these individual characteristics potentially biases the estimates of location effects.

To address possible bias due to omitting individual characteristics, Glaeser and Maré (2001) pioneered the use of panel data with individual fixed effects. Individual fixed effects will represent all the time-invariant differences in individual characteristics that are 'unobserved' from the data. Empirically, the estimate of the urban wage premium decreases significantly when individual fixed effects are introduced. In Glaeser and Maré (2001)'s study, workers in the US in densely-populated metropolitan areas and non-densely-populated metropolitan areas earned around 25% and 15% respectively more than workers in other areas when controlling for basic demographic and human capital variables. When individual fixed effects are introduced, these wage premium estimates decreased by around half. Yankow (2006), also using NLSY79, found that wage premiums, after controlling for race and experience, were around 20% and further decreased to 5% for big cities when individual fixed effects were introduced. The same patterns were found in the UK by D'Costa and Overman (2014), using the Annual Survey of Hours and Earnings data. The wage premiums for London, the UK's big cities and small cities were 23.5%, 6.2% and 4.8% respectively with OLS estimates, and decreased to 7.1%, 2.5% and 1.4% with Fixed Effects estimates. Even though controlling for individual heterogeneity reduces estimates of the urban wage premium by around half, studies still find that the relation between urban status and individual wages is positive and significant. While findings around the world show that workers in urban areas earn more than workers with similar skills in rural areas, little is known for Australian workers.

3.2.3 Urban learning and transitional noise

Even though individual fixed effects effectively deal with the time-invariant part of unobserved individual characteristics, they do not address the evolution of individual skills that is dependent on locations. Marshall (1890)'s hypothesis on the acceleration of human capital accumulation in cities suggests that a worker' skills appreciate with time that he spent in cities. A year in a large city potentially benefits wage growth more than a year in other areas through its effect on worker skills. Omitting the effects of experience that are specific to locations could bias the Fixed Effects estimates of the urban wage premium.

Glaeser and Maré (2001) found that workers who move from rural to urban areas experienced a wage rise of 7.9% within one year of migration relative to rural stayers. The rise is 11.1% for one to three years after migration and 11.8% for five or more years after migration (results with OLS estimates, using NLSY data). Yankow (2006) includes dummy variables in Fixed Effects models indicating whether an individual stays in cities, moves in cities or moves out of cities between years t and $t+1$ and estimates the models using First Difference. In that study, a year in a city associated with a 1.3% increase in individual wages relative to a year in a rural area. These results suggest urban learning benefits as rural-to-urban migrants' skills increase with time spent in urban areas, increasing the wage gap relative to rural stayers over time.

A more comprehensive study on urban learning and its effects on estimates of an area's wage premium was conducted by De la Roca and Puga (2017). Their strategy is to include years of work experience for each type of city, namely first and second biggest cities and third to fifth biggest cities, in individual wage equations, allowing the value of experience to differ based on different local economy sizes. A result that the value of experience is higher for larger cities will support the learning hypothesis. The findings are as expected; the first year of experience in first or second biggest city, Madrid or Barcelona, raises earnings by around 3.1% and the first year of experience in the third to fifth biggest cities raises earnings by around 1.6% relative to the same year in other cities. Furthermore, the estimate of wage elasticity with respect to city size is 0.024 in case that experience is specific to city types, not much different from the elasticity of 0.022 obtained when the specific effects of experience are ignored. The authors argue that the bias resulting from omitting urban learning benefits is potentially small, especially when migration flows among areas are balanced in datasets.

Wage premium estimates by First Difference are potentially less affected by urban learning benefits than Fixed Effects. The reason is that First Difference estimates the urban wage premium via the shift in individual wages immediately before and after the move while urban learning benefits are accumulated over time. However, there are several reasons why the wage shift may not accurately reflect an area's wage premium. Ashenfelter and Card (1984) observed a fall in the earnings of trainees relative to the comparison group before the trainees participated in training programs. An Ashenfelter's Dip would mean that wage gains from rural-to-urban migration will overestimate the urban wage premium and wage loss while moving from urban-to-rural will underestimate the premium. Another possibility is that migrants consider 'long run' labor market outcomes rather than immediate wage gains. If so, rural-urban migrants might receive less than the urban wage premium immediately after moving but experience high wage growth in subsequent years postmigration (Glaeser & Maré, 2001; Rowe et al., 2017; Yankow, 2006). Individual wages will follow similar patterns if migrants go through a settlement period, looking for suitable jobs or getting used to the new environment, when they earn less than their ability would suggest. Migrants' wages will recover in later periods and reflect their skills and the area's wage premium as normal. Both the relative wage dips immediately before and after migration will result in inaccurate First Difference estimates of the urban wage premium.

Similar to the pattern observed by Glaeser and Maré (2001) and Yankow (2006) for the US, Australian rural-to-urban migrants receive an immediate wage loss upon arrival but their wages recover in the subsequent years post migration, up to the wage levels of urban natives (Rowe et al., 2017). It is not clear from the studies of Glaeser and Maré (2001), Rowe et al. (2017) and Yankow (2006) whether high wage growth for rural-to-urban migrants in the years following migration indicates the high growth of individual skills in urban areas or simply reflects an acclimatization process where it takes time for individual wages to reflect their skills in a new environment.

3.2.4 The relation between individual wages and local economy size

Urban wage premium studies focusing on agglomeration economies are interested in further understanding the relationship between local economy size and local workers' income. While 'urban' and 'rural' indicate the size of a local economy, they are very broad groupings of different urban centers of different sizes. The existence of the urban wage premium indicates a positive relation between local economy size and local workers' income, but the estimates do not describe how local wages change when local economy size changes. One approach is to directly place a measure of local economy size such as employment density together with individual characteristics in a one-stage wage regression (Matano & Naticchioni, 2012). However, the more complete approach is the two-stage procedure suggested by Combes et al. (2008).

The benefits and technical details of the two-stage regression are in Combes et al. (2008) and Gobillon (2004). Using the first difference or time demeaning techniques to estimate a Fixed Effects model such as $y_{ict} = \mu_i + Z_{ct}\beta + \eta_{ct} + \epsilon_{it}$ introduces local error terms involving both η_{ct} and $\eta_{c'(t-1)}$ in the case of movers (Z_{ct} are location characteristic variables, including a measure of local economy size such as local employment density). Because of that complication, location variances cannot be estimated and nor can the standard errors of the coefficients (Gobillon, 2004). Ignoring the correlation of disturbances leads to biased standard errors according to Moulton (1990). In a two-step approach, individual shocks and location shocks are appropriately dealt with at each stage. All the unobserved location characteristics, including time-varying shocks, are contained in local-time fixed effects and can be dealt with more appropriately in the second stage (Combes & Gobillon, 2015).

In the first stage, location effects and individual effects are separated as in Glaeser and Maré (2001). Localization variables such as individuals' industry share of professionals and establishments are added in the first stage to control for local industry effects on wages. In the second stage, location effects estimated in the first stage are regressed on instrumented local density and other control variables. Controlling for location characteristics is necessary because location characteristics can affect both individual wages and local economy size. For example, areas close the coast potentially attract more people due to their friendly weather. Being close to the coast, and at the same time being economically accessible, increases local productivity and local workers' wages. Likewise, a university boosts local economic performance via its research outputs and increases the area's population via attracting students to the area. If so, the positive relation between local wages and local economy size does not imply agglomeration economies as expected. Furthermore, there is a potential feedback relationship between local economy performance and local economy size because people are potentially drawn to economically well-performing areas.

Related literature suggests the use of instruments in the second stage to address bias resulting from both omitting variables and the feedback relationship (Ciccone & Hall, 1996; Combes et al., 2008; Combes, Duranton, & Gobillon, 2011; Combes, Duranton, Gobillon, & Roux, 2010; De la Roca & Puga, 2017). An instrument for current local economy size needs to correlate strongly with the size for relevance and be uncorrelated with local economic performance for exogeneity. Historical population is a popular instrument choice, first used by Ciccone and Hall (1996) and later in other studies such as those of Combes et al. (2008, 2010), De la Roca and Puga (2017) and Matano and Naticchioni (2012). An area's past population predicts its current economic size because established population tends to last for a long time. It is exogenous because local factors that are attractive to people today and affect current local economy performance are different from factors that were attractive to people in the past. Another type of instruments is geological variables on soil characteristics (Combes et al., 2010; Rosenthal & Strange, 2008). The argument for characteristics of soil as instruments for current local economy size is like the argument for historical population. The quality of soil was important for agriculture-based economies in the past. Early settlements were drawn to areas with fertile lands, and the settlements still exist. The variables are exogenous because today's economy is less dependent on agriculture and human activities only have small impacts on soil characteristics (Combes & Gobillon, 2015).

Empirically, estimates of elasticity of wages with respect to density depends on the methods used, the zoning choices and especially individual skill controls. Combes et al. (2008), using a two-step procedure, found wage elasticity with respect to local density of around 3.5% for France. The elasticity estimates are consistent and statistically significant when OLS, FGLS, 2SLS, and First Difference estimates or instruments are used in second-stage regressions. De la Roca and Puga (2017), also using a two-stage regression, found that the static elasticity of wages with respect to city size was around 0.02, and medium-term elasticity for workers with around seven years of local experience was around 0.05 in Spain. More complete individual controls in the first stage in De la Roca and Puga (2017) probably explain why their estimates are smaller than those of Combes et al. (2008). Combes and Gobillon (2015) summarized related results and found that controlling for location characteristics as well as using instruments make small differences in the estimation of elasticities, decreasing the estimates by around 10% to 20%. On the contrary, controlling for individual heterogeneity by individual fixed effects has a large impact on estimates; the estimates decrease by more than half to typically around 2%.

3.2.5 Other concerns

Besides skills, estimation bias can result from the choice of the size and shape of discrete spatial units, known as the Modifiable Areal Unit Problem (MAUP). Briant, Combes and Lafourcade (2010), using French data, studied the effects of different zoning systems on spatial concentration and the estimation of agglomeration economies. The zoning systems in their study were grid (of equal squares), administrative, and 'partly random' zoning systems of different sizes. They concluded that, in general, variations resulting from estimation specifications are greater than those from the MAUP, and in the MAUP, size may be of higher concern than shape. They suggested researchers pay more attention to having the right specification for their research questions rather than to the MAUP.

One concern of models with individual fixed effects and the use of 'within-individual' variations is that the urban wage premium is identified via movers. That raises the question of self-selection bias where people who move are those who will gain most benefit from doing so (Combes et al., 2011; Gould, 2007). In such cases, gains in individual wages resulting from moving from rural to urban areas may simply reflect that movers are presented with better job opportunities rather than reflecting differences in productivity between locations. Glaeser and Maré (2001)'s results are suggestive of this selection process when urban-to-rural migrants did not receive an immediate wage loss. Their other result, that rural-to-urban migrants experienced a relative wage dip immediately before migration may at the same time indicate that workers who move to urban areas are those who were not particularly successful at their origin. Another concern is the correlation between moving from a rural to an urban area and entering a white-collar sector, which Gould (2007) found in the US, using NLSY data. If that correlation exists, the urban wage premium estimates may actually be wage gains from changing jobs or industries associated with rural-to-urban migration.

4 Theoretical model and econometric specifications

4.1 Theoretical model

The approach here is based on Combes et al. (2008) and Combes and Gobillon (2015), detailing how to derive $\ln(w_{ict}) = \ln(B_{ct}) + \ln(s_{it})$, as mentioned in the literature review. The derivation is described below.

For a representative firm located in employment area c in year t, the profit π_{ct} is given by

$$\pi_{ct} = p_{ct}Y_{ct} - \sum_{i\in(ct)} w_{ict}l_{ict} - r_{ct}z_{ct} \ (1)$$

where Y_{ct} and p_{ct} are the level of output and the output price, w_{ict} and l_{ict} are wages per hour and the working hours of worker i employed by the firm in employment area c in year t, z_{ct} and r_{ct} are the level of input and the input price. The firm decides the levels of l_{ict} and z_{ct} to maximize π_{ct}. Also assume a Cobb–Douglas production function

$$Y_{ct} = \frac{A_{ct}}{\alpha^{\alpha}(1-\alpha)^{1-\alpha}}\left(\sum_{i\in(ct)} s_{it}l_{ict}\right)^{\alpha} {z_{ct}}^{1-\alpha} \ (2)$$

A_{ct} is the local total factor productivity, s_{it} represents workers i's skills and the parameter $0 < \alpha < 1$. The firm chooses the level of labor l_{ict} and the level of input z_{ct} to maximize its profit given its profit function π_{ct} in (1) and production function Y_{ct} in (2). Substitute (2) into (1) and use the first-order conditions for firm profit maximization with respect to l_{ict} and z_{ct}

$$w_{ict} = A_{ct}p_{ct}\left(\frac{\alpha z_{ct}}{(1-\alpha)s_{it}l_{ict}}\right)^{1-\alpha} s_{it} \ (3)$$

$$r_{ct} = A_{ct}p_{ct}\left(\frac{\alpha z_{ct}}{(1-\alpha)s_{it}l_{ict}}\right)^{-\alpha} \ (4)$$

(3) and (4) yield the wages that worker i receives

$$w_{ict} = \left(\frac{A_{ct}p_{ct}}{{r_{ct}}^{1-\alpha}}\right)^{1/\alpha} s_{it} \equiv B_{ct}s_{it} \ (5)$$

or in log form

$$y_{ict} = \beta_{ct} + \ln(s_{it}) \ (6)$$

with $y_{ict} \equiv \ln(w_{ict})$ and $\beta_{ct} \equiv \ln(B_{ct})$

4.2 Econometric specifications

For simplicity, I also assume that location effects are relatively stable over the period or β_{ct} could be represented by β_c. The more comprehensive approach would be to allow location effects varying over time, accounting for location developments. Estimating each β_{ct} is demanding in terms of individual data. For C areas and T periods, there are CxT of β_{ct}. Using a Fixed Effects model, there should be both stayers and movers for each location at each period for β_{ct} to be identified. One location-time effect β_{ct} is set to zero and the others are identified through connectivity to the base. Thus, any pair of location times should be linked by worker movements in continuous periods. Combes et al. (2008)'s dataset includes 2,664,474 individual-year observations to estimate 341 employment areas over six periods. This study keeps β_{ct} time-invariant for data requirements and simplicity. Furthermore, De la Roca and Puga (2017) find that using fixed location effects with fixed city sizes lead to similar results to using time-varying location effects with time-varying city sizes.

4.2.1 Models without individual fixed effects and Fixed Effects models

From Equation (6), specific econometric models then depend on how individual skill part $\ln(s_{it})$ is accounted for. Assume that $\ln(s_{it})$ can be estimated by some available measures of individual characteristics X_{it} so that $\ln(s_{it}) = \mu_i + X_{it}\,\theta + \epsilon_{ict}$ where θ is the vector of parameters corresponding to the individual characteristics and ϵ_{ict} are random errors, Equation (6) becomes

$$y_{ict} = \beta_c + X_{it}\,\theta + \epsilon_{ict} \; (7)$$

Equation (7) can be estimated by OLS.

The benefit of not using individual fixed effects is that the effects of different skills on the urban wage premium are examined separately. In a Fixed Effect model, all time-invariant characteristics are pooled into individual fixed effects. Furthermore, OLS uses cross individual variations as well as within individual variations rather than only within-individual variations like Fixed Effects estimates or First Difference estimates for Fixed Effects models. Estimating wage premiums through movers assumes that stayers would receive the same benefits as movers if they moved. That may not always true because of the potential self-selection process (Combes et al., 2011; Gould, 2007, mentioned in the literature review).

Concerning unobserved innate individual abilities that can bias estimates of the urban wage premium, as noted by Glaeser and Maré (2001), we can add individual fixed effects μ_i to wage equations

$$y_{ict} = \beta_c + \mu_i + X_{it}\,\theta + \epsilon_{ict} \; (8)$$

μ_i represents other individual abilities that are not accounted for by X_{it} . That treatment is standard in labor economics. The benefits of Fixed Effects models in the case of the urban wage premium are discussed in greater details in Glaeser and Maré (2001), which I summarize in the literature review. Equation (8) can be estimated by Fixed Effects or First Difference, using within individual variations. β_c is estimated by including $C-1$ location indicators τ_{ict} which take the value of one if the person is in c year t and zero otherwise. The first and second analyses consider three sizes of urban centers: Major Urban, large urban centers, Other Urban, small urban centers, and Rural Area ($C=3$). More detailed descriptions of the structure are in the section on zoning structures.

In addition to individual characteristics, like Yankow (2006), I add union membership indicators and firm size indicators to X_{it} in some regressions. Union wage effects may be slightly positive, ranging from 2% to 11% (Cai & Liu, 2008; Cai & Waddoups, 2011; Waddoups, 2005) or even negative (Nahm, Dobbie, & MacMillan, 2017) in Australia. Earnings potentially increase with firm size (Brown & Medoff, 1989). Unions and large firms may be more active in urban areas than in rural areas. The two variables examine whether urban wage premiums, if they exist, are due to unions and large firms in big cities.

4.2.2 First Difference estimates

Assume that there are only two areas, a rural area and a city and all migrations are from a rural area to a city. The location effect of the city relative to the rural area $\beta_{City}=\delta$, and each year they work in the city, workers acquire skills that increase their wages by σ relative to working in rural areas. A city migrant will enjoy an extra wage of $\delta+n\sigma$ after n years after migrating to the city; the average wage benefit associated with cities estimated by the Fixed Effects estimation is $\hat{\delta}_{FE}=\delta+\frac{(n+1)}{2}\sigma$. In that case, the urban wage premium δ is overestimated by $\frac{(n+1)}{2}\sigma$. If all migrants were to move in the other direction from the city to rural area after n years in the city, the estimated wage benefits associated with living in the city is $\hat{\delta}_{FE}=\delta-\frac{(n+1)}{2}\sigma$. In that case, Fixed Effects underestimate the urban wage premium by $\frac{(n+1)}{2}\sigma$ (Combes et al., 2011; de la Roca & Puga, 2017). If movement among areas is balanced, $\hat{\delta}_{FE}$ will probably be close to δ as these biases cancel each other out.

The migration directions are slightly from urban to rural in the study sample (as shown in Table 3). The above argument and potential learning acceleration in urban areas mean that the Fixed Effects estimates may be biased downwards. Consider the first difference of Equation (8) between years $t+1$ and t

$$\Delta y_{ict}=\Delta X_{it}\theta+\beta_c\left(\tau_{ic(t+1)}-\tau_{ict}\right)+\Delta\epsilon_{ict} \ (9)$$

In this specification, β_c is identified when $\tau_{ic(t+1)} \neq \tau_{ict}$. Thus, β_c is identified via movers and is associated with wage changes right before and after the move rather than wage changes for the whole studied period like Fixed Effects. The longer time a person stays in an area or the larger n, the greater the area affects his skills via affecting his skill accumulation. Within a short timeframe of a year, a jump in wages associated with location changes, controlling for other skill changes, probably better reflects the difference in location effects between the locations of arrival and departure. In the first case when all migrations are from rural to city, $\hat{\delta}_{FD} = \delta + \frac{\sigma}{2}$ and in the second case when all migrations are from city to rural $\hat{\delta}_{FD} = \delta - \frac{\sigma}{2}$. Besides the advantage of First Difference estimators that they are less affected by learning-location time dependence, I use First Difference to check the robustness of the results by Fixed Effects.

With the three SOS areas in the study, Equation (9) can be written as

$$\Delta y_{ict} = \Delta X_{it}\theta + \beta_{Major}\Delta Major\ Urban + \beta_{Other}\Delta Other\ Urban + \Delta\epsilon_{ict} \quad (10)$$

Rural Area is the base. $\Delta Major\ Urban$ takes the value of one if worker i is outside Major Urban in year t and in Major Urban in year $t+1$, minus one if he moves out, and zero otherwise.[2] The same applies for $\Delta Other\ Urban$. To understand further whether changes in wages associated with the move reflect changes in location effects, following Yankow (2006), I augment $\Delta Major\ Urban$ into 'Move in Major Urban' and 'Move out Major Urban' indicators and do the same for $\Delta Other\ Urban$. The 'Move in' indicators take the value of one if people are not in the area in year t and in the area in year $t+1$ and the 'Move out' indicators are the reverse. The purpose is to allow the effects of moving in and out to be different rather than to constraint them to be the same as in Equation (9) and Equation (10). If wage change is solely associated with change in location effects, we expect that a worker who gains a wage premium when moving into an area will lose the same premium when moving out of the area. If so, the estimated coefficients of Move in and Move out will be equal in magnitude but have opposite signs.

Second, I add two indicators, Stay in Major Urban and Stay in Other Urban. The indicators take the value of one if the person stays in the same area in both years and zero otherwise. The specification examines whether people staying in Major Urban or Other Urban receive learning benefits relative to staying in Rural Area, as proposed by Glaeser (1999) and Marshall (1890). If urban learning benefits exist, everything else being equal, wage growth of urban stayers will be higher than wage growth of rural stayers or the

[2] $\Delta Major\ Urban$ takes the value of zero when the individual does not move or move between Other Urban and Rural Area between years t and $t+1$.

estimated coefficients of the two indicators will be positive. The reference group with this specification consists of people who stay in Rural Area in both years t and $t+1$.

As discussed in the literature review, the immediate wage shifts may not reflect the change in location effects due to possible transitional noise. I adapt Yankow (2006)'s long difference models to address that concern. Like before, $\Delta Major\ Urban$ and $\Delta Other\ Urban$ are augmented into 'Move in' indicators and 'Move Out' and 'Stay in' indicators are included. While these indicators take the value for the period between years t and $t+1$, the changes in individual wages as well as individual characteristics are for longer periods. For example, in a two-lagged difference model, changes in individual wages and individual characteristics are between years t and $t+2$. Similar set-ups are for longer lagged difference models; for example, for three-lagged difference, the wage change is for the period between years t and $t+3$, etc. As in the previous specification, the reference group consists of people who stay in Rural Area in both years t and $t+1$.

For long difference models, I also condition that individuals do not move in other years in the period. For example, in the two-lagged difference specification, people stay in the same place in years $t+1$ and $t+2$ and in the three-lagged difference, people stay in the same place from years $t+1$ to $t+3$. In this way, people who move to an area in year $t+1$ are in the same area in the later years in the period. The wage changes associated 'Move in' or 'Move out', controlling for individual skill development, will better reflect areas' wage premiums. Furthermore, as individual wages right before the move in year t might suffer an Ashenfelter's Dip as discussed, making wages in year t do not truly reflect individual ability, I extend the long difference models for wage changes for the period between years $t-1$ and $t+2$ (one year before the move to one year after the move) to the migration decision between years t and $t+1$.[3]

4.2.3 Two-step regression procedure

Rather than including both individual skills and local employment as right-hand side variables in a one-stage regression to examine the relationship between individual wages and the size of local economy, I follow Combes et al. (2008) and consider a two-stage regression

$$y_{ict} = \mu_i + X_{it}\theta + \beta_c + \epsilon_{ict} \quad (11)$$

$$\widehat{\beta_c} = Z_c\gamma + \eta_c \quad (12)$$

[3] As in the previous condition, the observations are from individuals who stay in the same place in year $t-1$ and t, and in the same place in year $t+1$ and $t+2$. Thus, people either change places between t and $t+1$ or stay in the same place for the whole period from year $t-1$ to $t+2$.

The literature review summarizes the advantages of a two-step procedure over a one-step procedure (see Combes et al. (2008) and Gobillon (2004) for more details). Individual characteristics are filtered out as before in the first stage in Equation (11). Rather than Major Urban, Other Urban and Rural Area, location effects are estimated for each labor market region in Australia. I exclude labor market regions with 'rural' characteristics and end up with 67 labor market regions for this analysis. I include the detailed restrictions in the results and discussion section.

In the second stage in Equation (12), the estimated location effects $\widehat{\beta_c}$ are obtained from Equation (11). Z_c and γ are labor market region c's characteristics and the corresponding vector of parameters. η_c are random errors at area level. Z_c includes local employment density in log form $\ln(emp\ den)_c$, a measure of the size of the local economy (Ciccone & Hall, 1996). Local employment better represents the size of the local market than local population as population includes people who do not work and not contribute directly to the local economy. The procedure measures the elasticity of individual wages with respect to employment density.

I check the robustness of the base results by comparing them to results from some other specifications. In one specification, the elasticity is estimated by placing $\ln(emp\ den)_{ct}$ directly into Equation (11) in a one-stage regression (Matano & Naticchioni, 2012). In another specification, individual fixed effects are not used in Equation (11) and $\widehat{\beta_c}$ is obtained by OLS. Lastly, urban economics theories suggest that location effects on wages are industry specific. Individual wages are affected not only by location characteristics but also by characteristics of local industries.[4] Following Combes et al. (2010), I include the terms $emp\ share_{cst}\delta_s$ in the first-stage regression. Here, $emp\ share_{cst} = \frac{emp_{cst}}{emp_{ct}}$ is a measure of local industry specialization where emp_{cst} is the number of workers employed in industry s in location c in year t, and δ_s is a parameter corresponding to industry s. Industry effects on wages, $emp\ share_{cst}\delta_s$, are allowed to be different among employment markets and the effects correspond to the industry's local specialization. By including the terms in the first-stage regression, industry location specific effects via specialization are filtered out. In that case, Equation (11) becomes

$$y_{icst} = \mu_i + X_{it}\theta + \beta_c + emp\ share_{cst}\delta_s + \epsilon_{icst}\ (13)$$

The estimates of the elasticities in the second stage could be biased due to omitted location characteristics as well as feedback relations between local economy size and local economic performance. I use

[4] Part of the industry specific effects is accounted for by including industry indicators in X_{it}. With the industry indicators, industry effects are assumed to be the same for different locations.

population in 1911, bulk density, available water capacity and terrain ruggedness as instruments for local employment density. The rationales for using historical population and soil characteristics as instruments for local economy size have been discussed in the literature review. Soils with high bulk density tend to hinder plants' root growth. Water and soils' ability to retain water, on the other hand, are crucial for plant development. The argument for terrain ruggedness is that flat terrains are able to support large population settlements (Combes et al., 2010). Weber (1899, p. 2) observed that environment factors such as climate and soil affected population distribution. Mountains had fewer inhabitants than valleys. The descriptions of these instruments are in the data descriptions section.

Like other economies, the Australian economy has experienced structural changes. In the early days after European settlement in 1788, the Australian economy relied mainly on wool exports and rural commodities exports to European markets. Economic developments in the 19th century were marked by major depressions and a gold rush following gold discoveries in Victoria in 1851. For the period from 1891 to 1973, Australian manufacturing experienced high growth and its output peaked in the 1960s, at about 28% of the country's outputs. More recently, Australia suffered from persistent inflation and high unemployment. Manufacturing sectors experienced decreases in outputs and the number of workers, and the Australian economy has become more service- and technology-oriented. (Attard, n.d.). Australian economic booms and busts, economic reforms, demographic changes, wars and the change from an agriculture-based economy to an industry and service-based economy give weight to the exogeneity of soil characteristics and historical population as instruments of current economy size.[5]

[5] Furthermore, as agriculture is not the main industry in the studied areas (areas with less than 10% of the workers employed in agriculture), soil characteristics in these areas potentially have small impacts on the area' current economic performance.

5 Zoning structures

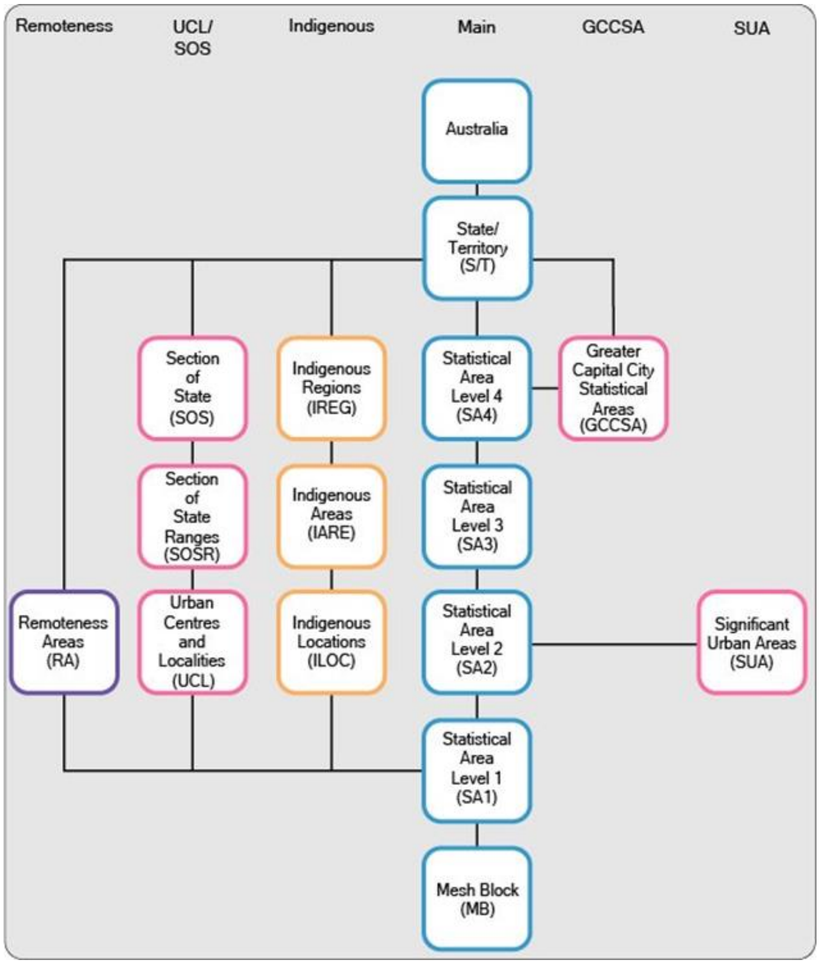
5.1.1 Geographic areas in Australia

From 2011, the Australian Statistical Geography Standard (ASGS) has provided a framework of statistical areas in Australia. The two main types of structures are the ABS structures and the non-ABS structures. The ABS structures are defined and maintained by the ABS for their output statistics (ABS, 2018b). The structures are hierarchical, in which a higher-level area is built from a cluster of lower-level areas. The non-ABS Structures are designed by other organizations; some examples of non-ABS Structures are local government areas, used by the government for administrative purposes, and postal areas, used by Australia Post for deliveries.

For the study, ABS structures offer several advantages over non-ABS structures:

- An ABS area is stable for five years (ABS, 2018b), allowing better comparisons of area data over time. Non-ABS structures are subject to more regular changes; for example, the government updates local government areas annually to match actual developments in the areas. Keeping track of and accounting for these changes over time is challenging in a panel study.
- Even though the ABS provides some statistics for non-ABS structures, it mainly provides data for ABS structures. Finding necessary information about non-ABS structures, on the other hand, can be challenging.

Figure 1 ASGS ABS Structures (ABS, 2018b)



5.1.2 Section of State (SOS) structure

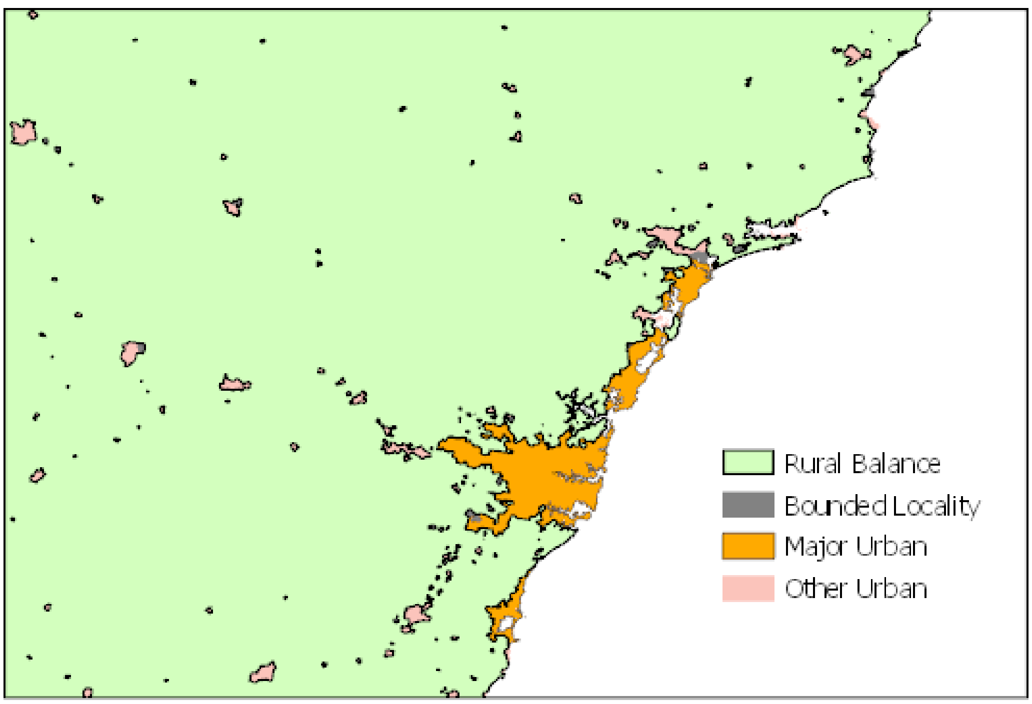
The SOS structure (edition 2011) divides Australia into four areas of Major Urban, Other Urban, Bounded Locality and Rural Balance, suitable for the study's purpose of examining rural-urban wage differentials. Rural and urban here follow the ABS's definition of Urban Centres and Localities. In essence, Urban Centres and Localities are areas with high population density or with 'urban' infrastructure. Urban infrastructure includes facilities such as airports, parks, education institutions, hospitals, office complexes, sport facilities and shopping centers. Features such as mines, wineries, dams, forests, national parks, etc., on the other hand, are not considered 'urban'. In more detail, an Urban Centre has population density of at least 200 persons per sq. km or a dwelling density of at least 50 dwellings per sq. km. A Locality, like an Urban Centre, presents urbanization, but at a smaller scale. The aggregate population of a Locality is at least 200 persons while that of an Urban Centre is at least 1,000 persons. (ABS, 2012b)

The SOS areas are aggregates of Urban Centres and Localities by their population.

- Major Urban represents all Urban Centres with a population of 100,000 or more.
- Other Urban represents all Urban Centres with a population ranging from 1,000 to 99,999.
- Bounded Locality represents all Localities.
- Rural Balance represents the remainder of Australia. (ABS, 2012b)

The three SOS areas, Major Urban, Other Urban and Locality, built from Urban Centres and Localities, representing cities and towns of different scales. The rest of Australia is Rural Balance having little population clustering and few urban activities. Even though Rural Balance covers more than 99% of the country's area (ABS, 2012c), in the study sample, around 70% of observations are in Major Urban while only around 2% and 8% are in Bounded Locality and Rural Balance respectively. Because not many persons are in Bounded Locality and Rural Balance, I merge these two areas and call them Rural Area in this study. Rural Area represents all small population clusters of below 1,000 persons.

Figure 2 Section of State, Sydney area (2011 edition)



Based on ABS cat. no. 1270.0.55.004 (ABS, 2012c)*. Major Urban covers areas around City of Sydney and along the coast, especially to the north of Sydney and other population centers are fragmented in Rural Balance.*

5.1.3 Main structures

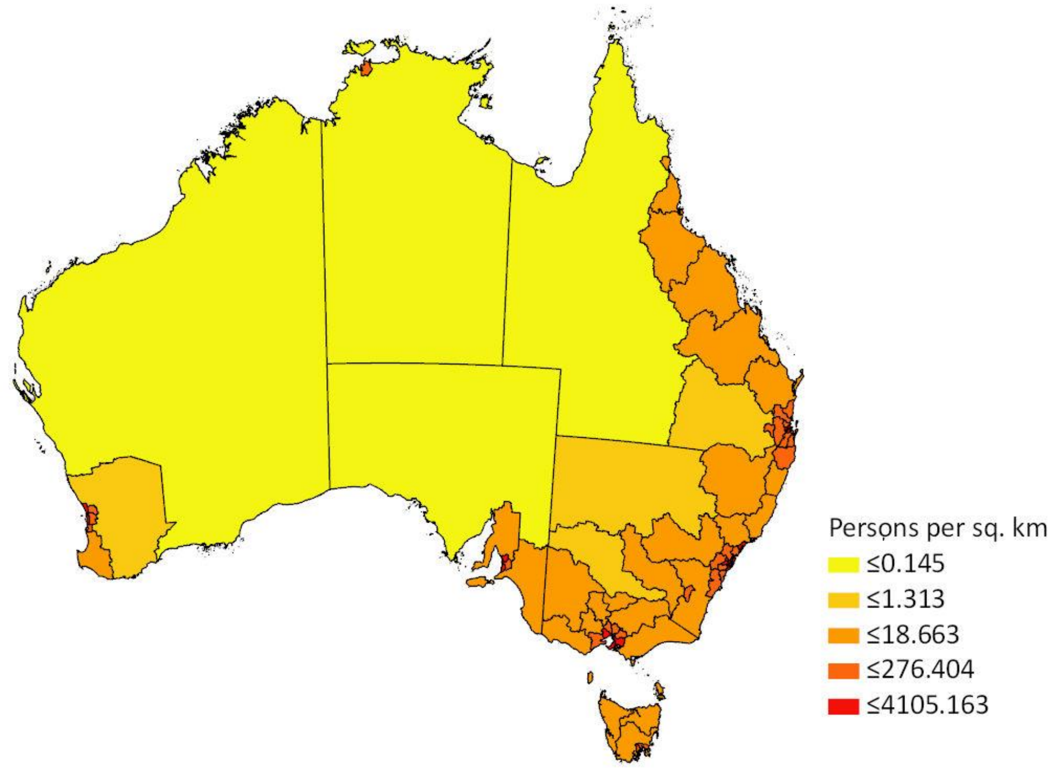
ABS cat. no. 1270.0.55.001 gives detailed description of the Main structures (ABS, 2016b). The labor market regions used in this study is Statistical Area 4 (SA4). As shown in Figure 1, a SA4 is built from whole SA3s and is below State/Territory level. In regional and remote areas, a SA4 contains similar and small local labor markets. In cities, a SA4 represents a major labor hub.

To attribute location effects to one area, it is desirable that that the area 'contains' people's daily activities and that there are few movements across the boundary. The ABS designs the Main structures based on the concept of a functional area, an area that many people commute and travel to, to access services within. A SA1 is 'internally connected by road' and a SA2 contains a service center and its functional area. A SA3 in a city area shares a transportation hub in city areas and in a regional and remote area has 'a distinct identity and similar social and economic characteristics' (ABS, 2016b). These criteria of Main structures are like the criteria for French employment areas used in Combes et al. (2008) where area boundaries are defined based on commuting patterns. Another advantage of the Main structures is that locations within an area are likely to share similar characteristics. As a result, an area's statistics are representative and do not hide large variations within. This applies to location variables in the study such as local employment density.

The study uses labor market regions because they are the main areas for the ABS's employment data output. Another advantage is that both labor market and labor supply locations are contained in a region. Ideally, the location where an individual works directly affects his work performance and wages, and its effects rather than the residence location's effects should be measured. Unfortunately, the HILDA survey only has questions on workplace postcodes and suburbs from wave 17 onwards. Using labor market regions increases the chance that a residential area is also the workplace area.

There are 106 SA4s (edition 2011) in Australia including 18 non-spatial areas for special purposes (migratory – offshore – shipping and no usual address). There are relatively large variations in both in area and population among the 88 spatial SA4s. The four largest areas, namely Western Australia – Outback, Northern Territory – Outback, Queensland – Outback and South Australia – Outback, cover around 75% of the whole of Australia. Sydney – Eastern Suburbs has the smallest area, only around 58 sq. km (ABS, 2010). Figure 2 shows that most Australian population centers are along the East and South East coasts. The four outback areas only have population density (based on population figures in 2011 from the ABS) of below 0.145 persons per sq. km while areas close to Australian capital cities are much smaller and have much higher population density. As of 2016, the three most populous areas are around the City of Melbourne: Melbourne – South East, 793,612 persons, Melbourne – West 765,986 persons, and Melbourne – Inner, 635,933 persons (ABS, 2018a).

Figure 3 Statistical Area Level 4 (2011 edition)



By ArcGIS, based on the ABS's SA4 digital boundaries (edition 2011) (ABS, 2010) and the ABS's population grid in 2011 (ABS, 2014b).

6 Data descriptions

6.1 The HILDA survey

The study uses the Household Income and Labour Dynamics (HILDA) survey of Australia, Release 16 from 2001 to 2016. The HILDA is a longitudinal survey in Australia, focusing on family formation, income and work. The survey follows a nationally representative sample of Australian households and household members across years. The initial sample in 2001 had 7,682 responding households with 19,914 household members. Wave 16 in 2016 had 9,750 responding households with 23,496 household members including the top-up from wave 11 (Summerfield et al., 2017, p. 15). The HILDA survey contains rich information on individual income, individual characteristics and individual background for the study's wage equations. Individual residential locations by SOS structure is in the HILDA general release data while locations by labor market regions are only in the restricted release. In total, there are 40,746 individuals and 651,936 individual-year observations in the dataset.

6.1.1 Individual hourly wages

I calculate nominal wages, w_{it}, following the HILDA's guide in Summerfield et al. (2017, p.63). The calculation is restricted to employed persons who report earning current or usual wages or salaries and who usually work a positive number of hours per week. Employed persons are persons who are aged 15 years old or over and participated in economic activities during the week before the interview. The definition of employment includes employee, employer/self-employed, employee of own business and unpaid family workers as being employed. I exclude unpaid family workers and employer/self-employed because these individuals do not often receive and report wages and salaries, making the reported hourly wage calculations for these two groups unreliable. Following the labor economics tradition discussed in the literature review, I focus on full-time male workers, working more than 35 hours per week (in all jobs). The usual weekly gross wages and salaries (before tax) in all jobs (imputed) include earnings from main jobs and from other jobs. Hourly wages are the total gross earnings divided by the combined weekly hours that the person usually works across all jobs (Summerfield et al., 2017, p. 63). With these restrictions, the study sample consists of 9,257 individuals and 54,636 individual-year observations with valid location information.

6.1.2 Individual characteristic variables

Years of schooling can be estimated based on the highest level of education achieved, such as in Peng Yu (2004) and Leigh and Ryan (2005). The number of years are given depending on how long it takes typically to achieve the education level, for example, 18 years for a master's degree or doctorate, 16 for a bachelor's degree, 14 for advanced diploma or diploma, etc. (Peng Yu, 2004). This calculation may ignore the fact that people take different paths to the highest level and that the qualifications obtained may not all be relevant in achieving the highest level. Two people may reach the same highest level of Doctoral Degree, but one can have more qualifications than the other. In this study, I try a different approach to better measure an individual's total effort in schooling activities. The HILDA has variables on the highest year of school completed and the number of qualifications with Australian Standard Classification of Education (ASCED) codes, obtained after leaving school to the time of the interview. I match each qualification with its notional duration suggested by the ASCED and the Australian Qualifications Framework 2013 (AQF), using its ASCED code. The notional duration expressed in full-time years is the time needed to achieve the course's learning outcomes (Australian Qualification Framework Council, 2013).[6] Years of schooling is the sum of years of school education (the highest year of school completed plus one year for Kindergarten/Preparatory) and the total estimated duration of all qualifications obtained after leaving school. Extreme values of more than 30 years of schooling are excluded. The possible drawback of this 'years of schooling' is that it may be prone to error as respondents may forget or mistakenly report the type and number of qualifications they have.

Years of work experience is accumulated time in paid work (years).[7] The HILDA only calculates the variable for individuals who report their work time in subsequent years from the first interview by adding work experience reported in the first interview to each subsequent period's work experience. Year 12, Bachelor (or Honours) and Postgraduate (Masters or Doctorate) variables indicate the highest education achievements. Postgraduate takes the value of one for Masters' or Doctoral Degree owners. Bachelor takes the value of one for the highest education qualification of Bachelor (or Honours), and Year 12 is one for completing Year 12 but not having a bachelor's degree. Two additional work experience variables in the HILDA are years in current occupation and years with current employer. Time in current occupation will predict the amount of occupation specific skill. Likewise, spending more time with the same employer can increase performance as workers get used to the staff and procedures.

In addition to the traditional Mincerian measures of human capital, the HILDA data contains measures of individual cognitive ability and personality. These variables provide direct measures of individual skills, in the spirit of AFQT scores in the NLSY data and Bacolod et al. (2009)'s motor and cognitive skills. The three cognitive tests are Backwards Digit Span, Simple Digit Modalities and National Adult Reading Test (short

[6] The durations for different type of qualifications used in the study are provided in the appendix.
[7] Rather than that, years of work experience can be measured via age and years of schooling, $age - years\ of\ schooling - 5$ (Peng Yu, 2004). In Australia, States and Territories are responsible for school education. The age that children start compulsory education is around 5 to 6 years old, varying slightly between areas.

form or NART25). The Backwards Digit Span test measures working memory span where interviewees are asked to repeat given strings in reverse order. The Simple Digit Modalities test is used to detect cerebral dysfunction or to measure 'divided attention, visual scanning and motor speed' in general (Strauss, Sherman, & Spreen, 2006, p. 617). NART25, a reading test of 25 irregularly spelled words, is a short form of the National Adult Reading Test. As reading ability highly correlates with intelligence, the NART25 score also indicates individual intelligence (Wooden, 2013). The HILDA provides scores on the Big Five Personality Traits, namely extraversion, emotional stability, openness to experience, agreeableness and conscientiousness. The higher a participant scores on a personality trait, the more the trait describes the person (Summerfield et al., 2017). Cognitive tests were in waves 12 and 16, and personality tests were in waves 5, 9 and 13 only. Because both cognitive ability and personality are relatively stable for working-age adults (Cobb-Clark & Schurer, 2012) and to preserve observations in regressions, I use an individual's average score on a type of test as his score on that test in all waves.

Occupation indicators in the study follow the major occupation groups of the Australian and New Zealand Standard Classification of Occupations 2006 (1-digit ANZSCO 2006). There are eight ANZSCO 2006 major groups: managers, professionals, technicians and trade workers, community and personal service workers, clerical and administrative workers, sales workers, machinery operators and drivers, and laborers. The ANZSCO groups are linked with skill levels that workers in the group often possess. The ABS describes the skill level as 'a function of the range and complexity of the set of tasks performed in a particular occupation' (ABS, 2006, p. 6). In the sample, Technicians and Trade has the most observations, at 23%. Major Urban has higher percentage of Professionals, at 25%, compared to Other Urban and Rural at 13%.

Industry indicators follow major groups in the Australian and New Zealand Standard Industrial Classification 2006, Revision 2.0 (ANZSIC 2006 division). There are 19 industrial groups in the classification. In the sample, Major Urban's workers are more likely to be in the 'Professional, Scientific and Technical Services' industry or the 'Public Administration and Safety' industry while Rural Area's workers are more likely to be in the 'Agriculture, Forestry and Fishing' industry. Industry indicators and occupation indicators control for individual differences in industry and occupation aptitudes. It is also true that including the variables in wage regressions accounts for industry effects on individual wages.[8]

[8] Individuals consider expected earnings when they choose their careers and plan their education accordingly. Thus, occupation and education variables can be endogenous. This study ignores these endogeneity issues for simplicity.

Overall, skill variables in the study consist of the traditional Mincerian measures of education and work experience, direct measures of cognitive ability and personality traits, and indirect measures of skills through occupation and industry (Bacolod et al., 2009). Other individual characteristic variables are married (or in de facto relationship) and born overseas. Age is not included because it is highly related to years of schooling and years of experience.

The union membership variable indicates whether workers are in a union or an employee association. The firm size variable in the HILDA takes values from one to seven corresponding to the number of the firm's employees nationwide: 1 for less than 20 persons, 2 for 20 to 99, 3 for 100 to 499, 4 for 500 to 999, 5 for 1,000 to 4,999, 6 for 5000 to 19,999, and 7 for 20,000 or more. In this study, I use firm size indicators for the size of the firm that an individual work for.[9] In the sample, urban firms are larger than rural firms while there is little difference among the three areas on the percentage of people in unions.

Table 1 shows that most of the observations are in Major Urban and Other Urban, 37,709 and 11,159 respectively. In the sample, the average wages are AUD 31.21 for Major Urban, AUD 26.53 for Other Urban and AUD 25.75 for Rural Area. The sample's average wages in Major Urban are around 21% more than in Rural Area. That premium for Major Urban is close to the premium of 17% to 22% for Australian capital cities reported by the Australian Treasury (2017, based on ABS data). Regarding skills across the three areas, Major Urban has a higher level of human capital than the other two areas. In the sample, Major Urban residents spend more years in school, have higher qualifications and score higher on cognitive tests (except the short NART). They are more likely to be born overseas. Regarding personality traits, people in Major Urban areas are more extroverted, conscientious and open to experiences but less emotionally stable than people in other areas. In the sample, workers in Major Urban spend more time studying and are slightly younger, and so they have fewer years of work experience in than workers from the other areas. In the sample, the average ages in Major Urban, Other Urban and Rural Area are 38.8, 38.4 and 40.5 years respectively.

[9] Responses on firm size of 'Don't know but fewer than 100' and 'Don't know but 100 or more' are not considered. There are seven firm size indicators corresponding to the HILDA's seven firm size groups.

Table 1 Sample statistics

	Major Urban		Other Urban		Rural Area	
	Mean	SD	Mean	SD	Mean	SD
Hourly wage	31.205	18.784	26.525	15.089	25.751	15.004
Hourly wage adjusted by inflation[10]	24.707	14.090	21.031	11.111	20.410	11.301
Married (or de facto)	0.716	0.451	0.721	0.449	0.735	0.442
Born overseas	0.250	0.433	0.088	0.284	0.114	0.318
Years of schooling	14.062	3.126	12.820	2.950	12.971	3.011
Years of experience	19.955	12.304	20.718	12.674	22.605	13.003
Postgraduate	0.072	0.259	0.020	0.141	0.025	0.158
Bachelor	0.250	0.433	0.116	0.320	0.141	0.349
Year 12	0.525	0.499	0.585	0.493	0.565	0.496
Years in current occupation	9.636	9.771	9.798	9.899	12.089	11.702
Years with current employer	7.114	8.061	7.312	8.471	8.379	9.499
Backwards Digit Span	5.130	1.341	4.924	1.255	4.907	1.302
Symbol Digit Modalities	50.927	9.946	48.639	10.407	47.129	9.633
Short NART	14.668	5.149	12.748	5.257	12.992	5.466
Extroversion	4.341	0.964	4.336	0.921	4.310	0.887
Agreeableness	5.161	0.797	5.110	0.830	5.109	0.810
Conscientiousness	5.039	0.900	4.961	0.886	4.982	0.886
Emotional stability	5.112	0.929	5.098	0.930	5.156	0.920
Openness to new experience	4.354	0.918	4.145	0.909	4.095	0.924
Union	0.264	0.441	0.322	0.467	0.277	0.447
Firm size	4.598	1.810	4.522	1.831	4.355	1.921
Observations of dependent variables	37,709		11,159		5,768	

6.2 Location characteristic variables

Location characteristic variables are used in the third analysis. The main variable is local employment density for Australian labor market regions.

$$emp\ den_{ct} = \frac{emp_{ct}}{area_c} \ (14)$$

[10] The study does not use inflation-adjusted wages as time indicators already account for overall changes in price levels. Wage adjusted by inflation in Table 1 is only for reference. The ABS's consumer price index covers eight State and Territory capital cities in Australia. The consumer price index used is from ABS cat. no. 6401.0. The numbers are for June quarter, all groups CPI (ABS, 2018c).

$emp\ den_{ct}$ (persons per sq. km) is employment density of labor market region c in year t. emp_{ct} here is measured by the employed total (full-time and part-time) in June of the year (ABS, 2018d). For the period from 2001 to 2016, all areas, except for areas around Australian capital cities, have employment density of below 500 workers per sq. km. All three areas with more than 2,000 workers per sq. km are around the City of Sydney: Sydney – City and Inner South, Eastern Suburbs and Inner West.

For a measure of local specialization, local industry share $emp\ share_{cst} = \frac{emp_{cst}}{emp_{ct}}$, I obtain emp_{cst} for each industry in each labor market region and emp_{ct} for each labor market region for the years 2011 to 2016 (ABS, 2018d).[11] Average employed total of the four quarters preceding August is taken as the employment figure for the year. Other derivations from ABS's data are explained in the related footnotes.

Australian historical population by local government areas is available from 1911 in ABS cat. no. 3105.0.65.001 (ABS, 2014a). Past local government areas have gone through amalgamations and name changes, and they are not comparable to today's local government areas. To link past local government areas' population figures to today's labor market regions, I purchased an Australian town list from AustralianTownsList.com (the data is based on ABS data). The list contains over 15,000 town name records with no duplicates, and their corresponding post codes. I matched the historical local government names to town names in the list and their current post codes. The post codes were then linked with labor market regions using the ABS's postcode-to-SA4 correspondence (ABS, 2012a). In most cases when historical names do not match, the area has been abandoned, and there are no other ways but to manually look for records to attribute population to the correct region. Overall, this is not an exact process. Some past local government areas lie on several labor market regions (especially near capital cities where the regions are small). However, a historical local government area is often within a labor market region because the regions are much larger areas. Furthermore, its name is often the name of its center town where most of the area's population concentrated. In most of cases, the town name is still used today and so available in the list. Therefore, the process likely attributes a large part of the past population to the right regions.

Other location characteristic variables are derived with the help of GIS software. Distance to coast (100 km) is the closest distance from a labor market region to the coast or a harbor. The measure is based on the ABS's SA4 digital map (ABS, 2010) and Australian digital boundaries (ABS, 2016a). Most of the regions in this study are adjacent to the coast and their distances to coast are zero. For instrument variables, used

[11] I used the catalogue issues in May 2018 and Feb 2015 of ABS cat. no. 6291.0.55.001. There are some minor differences between the two issues.

in the second stage of Combes et al. (2008)'s procedure, I overlay SA4's digital map with related gridded geographic data and find necessary statistics for each region. Many soil and landscape grids for Australia can be accessed via CSIRO's data access portal.[12] This study uses bulk density (whole earth) (g/cm^3) and available water capacity (%) by Viscarra Rossel, Chen, Grundy, Searle and Clifford (2014). A labor market region's water available capacity and bulk density are the means of these measures within the region. Similarly, terrain ruggedness at a fine scale for a region is measured by the standard deviation of all elevation values within the region.[13] The study uses water available capacity for top soils, 0–5 cm depth, and bulk density for soils is 15–30 cm depth. I include more detailed descriptions of the two measures and the values of the location characteristic variables in the data appendix.

[12] The website for Soil and Landscape Grid of Australia is http://www.clw.csiro.au/aclep/soilandlandscapegrid/ProductDetails-SoilAttributes.html

[13] Elevation values are from GEODATA 9 Second DEM (DEM-9s) Version 3 (DEM: Digital Elevation Model) (Geoscience Australia, 2008).

7 Results and discussion

7.1 Skills and urban wage premium

7.1.1 OLS estimates

To see how various skills affect urban wage premiums, I first ignore the possible unobserved individual heterogeneity and estimate the following relationship using OLS

$$y_{ict} = \beta_c + X_{it}\theta + \epsilon_{ict} \text{ (Equation (7))}$$

All variables are as defined in the section of econometric specifications. y_{ict} is the log of individual hourly wages, X_{it} and θ are individual characteristics and their corresponding coefficients, and ϵ_{ict} are random errors. β_c, the wage premium of location c (or location effects of c), is the parameter of interest. They are estimated by including $C-1$ location dummies, τ_{ict}; τ_{ict} takes the value of one if worker i is c at year t and zero otherwise. There are three areas in this analysis, Major Urban, Other Urban and Rural ($C = 3$). Rural Area is the base. All regressions include yearly time indicators to control for countrywide time shocks such as increases in overall price levels.

Individual characteristic variables X_{it} consist of marital status, born overseas, years of schooling, years of work experience and its square, education levels, three cognitive ability measures, five personality traits measures, occupation and industry indicators. I add them by groups to examine the effects of each type of skills on the urban wage premium.

Columns (1) to (7) in Table 2 present OLS results. Column (1) is the baseline regression with no controls for differences in individual skills. Workers in Major Urban, population centers from 100,000 persons, earn a wage premium of 19.3% relative to their counterparts in Rural. Workers in Other Urban, population centers of above 1,000 but below 100,000 persons, also earn a wage premium, albeit only about a fifth of Major Urban's premium, of 4.9%. The results are statistically significant at the 1% level. The premiums are less than the premiums of 30% often observed in other countries or the premiums of as much as 60% of Paris or Madrid relative to those countries' rural areas (Combes et al., 2008; De la Roca & Puga, 2017). Controlling for individuals' demographic differences in marital status and being born overseas in Column (2) does not change urban wage premiums significantly.

Columns (3) and (4) add traditional human capital controls of education and work experience. Column (3) considers a traditional Mincerian human capital specification with years of schooling, years of work experience and work experience squared. Column (4) provides more details on education and work experience, namely education levels and experience in current occupation and with current employer.[14] Education levels account for nonlinear effects of education on wages. I further include ANZSCO 2006 occupation and ANZSIC 2006 industry indicators in Column (4). Occupation indicates worker skills due to a market clearing process as shown by Bacolod et al. (2009). Wages vary by industries, and highly paid industries may be concentrated in large cities, explaining the urban wage premium. Accounting for individual differences in education, work experience, occupation and industry in Column (4) reduces Major Urban's premium by nearly half of the baseline estimate to 12.1%. Interestingly, Other Urban's wage premium increases slightly to 6.3%. The estimates are in line with the estimates obtained with similar specifications in Yankow (2006), 18.7% for Big City and 8.2% for Small City. In Yankow (2006), Big City has a population of more than one million and Small City has a population of more than 250,000 but less than one million, and that may explain why the estimates obtained are slightly higher. With the individual controls, the model now explains 43.2% of individual wage variations compared to only 15% of the variations explained by the model with no individual controls in Column (1).[15]

Columns (5) and (6) add direct measures of skills available in addition to Mincerian skill controls. Like Bacolod et al. (2009), Column (5) examines how individual cognitive ability affects estimates of the urban wage premium. Column (6) further adds the HILDA measures of the Big Five personality traits. Compared to traditional measures of skills, these direct measures of skills appear to have less impact on the estimations. The results are similar to those of Glaeser and Maré (2001) and Yankow (2006) when they found that a basic ability test, AFQT scores, made little differences in wage premium estimates. Among the personality traits, only conscientiousness and agreeableness are significant at the 1% level. While male workers are rewarded for being conscientious, they are penalized for being agreeable.[16]

Column (7) examines the extent to which urban wage premiums are due to union activities or firm size in urban areas. Controlling for these labor market differences reduces wage premiums substantially to 8.1%

[14] The reference group for education levels is 'below level 12'.

[15] Dobbie, MacMillan and Watson (2014), also using the HILDA survey, examined the wage effects of general experience (years of experience), occupational tenure (years in current occupation) and job tenure (years with current employer). They concluded that among the three, general experience is the most important determinant of individual wages. Job tenure effects, on the hand, disappear once unobserved individual ability and occupational tenure are accounted for. Table 2 shows results in line with Dobbie et al. (2014)'s conclusions; the point estimates on years with current employer are not different from zero in all regressions (except in Column (5)). The estimate of years in current occupation becomes zero when firm size is added in Column (11). It is possible that workers with long occupational tenure tend to work for large firms, and the positive relation between individual wages and occupational tenure is actually the large-firm wage premium (Brown & Medoff, 1989).

[16] The impacts of personality on earnings are well-researched. A study of Mueller and Erik Plug (2006) suggested that for male workers, conscientiousness did not generate earning returns while 'non-agreeableness' did.

for Major Urban. The point estimate on Other Urban is 2.3% and not statistically different from zero. Either union or firm size is important in explaining wage premiums in Australia; I later examine the two channels in more details with Fixed Effects models in the part following (in Columns (11) and (12) in Table 2).

Overall, OLS estimates show that earnings are positively associated with urbanization, a pattern found around the world. Considering individual heterogeneity with all the controls significantly reduces urban premiums; however, workers in Major Urban in Australia still earn 11.8% more than their counterparts in Rural Area. Interestingly, wage premium estimates for the smaller population centers, Other Urban, are not affected as much by individual controls, staying at around 5% during the process. The magnitudes of the premiums are largely in line with the literature (Yankow, 2006). D'Costa and Overman (2014), also with age, occupation and industry controls, obtained smaller estimates for the UK: the wage premium of 6.2% for big cities, and 4.8% for small cities in the UK.[17] Education and work experience are more important than other skills in explaining spatial wage differentials in Australia, similar to findings in Glaeser and Maré (2001) and Yankow (2006).

Personality traits, at least the Big Five traits measured by the HILDA, makes little difference in the wage premium estimates. That casts doubt on Combes and Gobillon (2015)'s suggestion that personality traits are among unobserved individual differences that if omitted, will bias wage premium estimates. Ambitiousness or adventurousness might not drive the bulk of spatial wage differentials as some may expect. Likewise, cognitive test scores, rough measures of individual cognitive speeds (Symbol Digit Modalities), intelligence (NART) and memory (Backwards Digit Span) have little impact on the estimates. Using individual fixed effects significantly reduce urban wage premium estimates as Column (8) shows, and there will be other differences in individual characteristics that are still unobserved. As personality traits and cognitive ability are not among these unobserved characteristics, the question is what these characteristics are?

[17] In D'Costa and Overman (2014)'s study, small cities represent areas with 100,000 to 250,000 persons and big cities represents areas with 250,000 to one million persons.

Table 2 Estimation of wage premiums considering individuals heterogeneity

	OLS (1)	OLS (2)	OLS (3)	OLS (4)	OLS (5)	OLS (6)	OLS (7)	Fixed (8)	Fixed (9)	Fixed (10)	Fixed (11)
Major Urban	0.193[a]	0.193[a]	0.167[a]	0.121[a]	0.116[a]	0.118[a]	0.081[a]	0.075[a]	0.077[a]	0.077[a]	0.046[a]
	(0.02)	(0.02)	(0.02)	(0.01)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)
Other Urban	0.049[a]	0.054[a]	0.073[a]	0.063[a]	0.048[a]	0.048[a]	0.023	0.046[a]	0.041[a]	0.041[a]	0.025
	(0.02)	(0.02)	(0.02)	(0.01)	(0.02)	(0.02)	(0.02)	(0.01)	(0.01)	(0.01)	(0.02)
Married or de facto		0.275[a]	0.120[a]	0.088[a]	0.086[a]	0.081[a]	0.065[a]	0.046[a]	0.045[a]	0.045[a]	0.033[a]
		(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)
Born overseas		0.039[a]	−0.019	−0.016	−0.007	−0.006	−0.001				
		(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)				
Years of schooling			0.045[a]	0.008[a]	0.003	0.003	0.001	0.027[a]	0.023[a]	0.023[a]	0.020[a]
			(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
Years of experience			0.035[a]	0.030[a]	0.031[a]	0.030[a]	0.028[a]	0.050[a]	0.048[a]	0.047[a]	0.049[a]
			(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.01)	(0.01)	(0.01)	(0.01)
Years of experience sq.			−0.001[a]	−0.001[a]	−0.001[a]	−0.001[a]	0.000[a]	−0.001[a]	−0.001[a]	−0.001[a]	−0.001[a]
			(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
Postgraduate				0.329[a]	0.327[a]	0.333[a]	0.330[a]		0.090[c]	0.091[c]	0.061
				(0.03)	(0.03)	(0.03)	(0.03)		(0.05)	(0.05)	(0.05)
Bachelor				0.241[a]	0.229[a]	0.232[a]	0.234[a]		0.056	0.057	0.003
				(0.02)	(0.02)	(0.02)	(0.02)		(0.04)	(0.04)	(0.04)
Level 12				0.095[a]	0.079[a]	0.078[a]	0.072[a]		0.084[a]	0.086[a]	0.003
				(0.01)	(0.01)	(0.01)	(0.02)		(0.02)	(0.02)	(0.03)
Years in current occupation				0.003[a]	0.003[a]	0.003[a]	0.003[a]		0.001[b]	0.001[b]	0.000
				(0.00)	(0.00)	(0.00)	(0.00)		(0.00)	(0.00)	(0.00)
Years with current employer				0.001	0.001[c]	0.001	0.000		0.000	0.000	0.000
				(0.00)	(0.00)	(0.00)	(0.00)		(0.00)	(0.00)	(0.00)
Backwards Digit Span					0.004	0.004	0.003				
					(0.00)	(0.00)	(0.00)				
Symbol Digit Modalities					0.003[a]	0.003[a]	0.003[a]				
					(0.00)	(0.00)	(0.00)				
Sort NART					0.005[a]	0.006[a]	0.007[a]				
					(0.00)	(0.00)	(0.00)				
Extroversion						0.008	0.012[b]				
						(0.01)	(0.01)				
Agreeableness						−0.022[a]	−0.025[a]				
						(0.01)	(0.01)				
Conscientiousness						0.035[a]	0.029[a]				
						(0.01)	(0.01)				
Emotional stability						−0.011[c]	0.005				
						(0.01)	(0.01)				
Openness to experience						−0.012[c]	−0.004				
						(0.01)	(0.01)				
Union							0.046[a]			0.035[a]	0.027[a]
							(0.01)			(0.01)	(0.01)
Firm size indicator							Yes				Yes
Occupation indicator	No	No	No	Yes	Yes	Yes	Yes	No	Yes	Yes	Yes
Industry indicator	No	No	No	Yes	Yes	Yes	Yes	No	Yes	Yes	Yes
N*	54,636	54,613	53,622	53,079	42,246	40,155	25,857	53,630	53,087	53,076	33,222
R sq.	0.150	0.203	0.328	0.432	0.436	0.441	0.487	0.381	0.389	0.390	0.441

All regressions include a constant term and time indicators. Dependent variable is log of nominal hourly wages. Numbers in brackets are robust standard errors, clustered on individuals. [a], [b] and [c] are significant at 1%, 5% and 10% respectively. Reported R sq. is overall for OLS regressions and within-individual for Fixed Effects regressions.

*Sample sizes are different because variables have different numbers of observations in the sample.

7.1.2 Fixed Effects estimates

Following Glaeser and Maré (2001), I assume the following relationship

$$y_{it} = \beta_c + \mu_i + X_{it}\,\theta + \epsilon_{it} \text{ (Equation (8))}$$

and estimate by Fixed Effects. β_c is identified through movers as for non-movers, τ_{ict} is fixed during the period and $\tau_{ict} - \bar{\tau}_{ic} = 0$. Table 3 shows that there are enough migrations in the sample, allowing reliable estimation of location effects.[18] Migrations for each area are relatively balanced between moving in and out, slightly towards urban to rural. Numbers of migrations in a direction for an area are around 400 in the sample.

Table 3 Migrations for Major Urban, Other Urban and Rural Area (Rural Area consists of Bounded Locality and Rural Balance)

	Move in	Move out	Stay
Major Urban	423	479	16,927
Other Urban	501	517	7,952
Rural	434	362	3,904
Total	1,358	1,358	28,783

The results are from Columns (8) to (11) in Table 2. Column (8) includes married, years of schooling, years of experience and its square. The regression has the coefficient on Major Urban of 0.075. Workers moving from Rural to Major Urban experience an 7.5% increase in wages relative to stayers. The coefficient is less than half of the OLS estimate with similar controls for individual differences in Column (3). The estimate is also significantly smaller than the OLS estimate with all individual controls in Column (6). The point estimate for Other Urban decreases, to a lesser extent, to 0.046. For movers from Other Urban to Major Urban, the predicted increase in wages is around 2.9% (i.e. $\approx 7.5\% - 4.6\%$).

Workers may move to cities to participate in white-collar work (Gould, 2007), and so any increase in wages is due to sectoral change rather than changes in location effects. Without controlling for industry and occupation, the endogenous location choice may bias the wage premium estimation. In this case, adding more detailed variables on education and work experience, occupation skills and industry indicators in Column (9) does not change the estimate of the wage premium by much. These estimates with Fixed Effects are in line with the findings of Glaeser and Maré (2001) and Yankow (2006) of around 5% to 10% for large cities and 2% to 7% for small cities.[19] Even though a broad range of individual characteristics,

[18] Migrations here are changes in location of individuals between two continuous years in the sample.
[19] Small cities are metropolitan areas with fewer than 500,000 inhabitants and large cities are metropolitan areas with more than 500,000 inhabitants in Glaeser and Maré (2001).

including cognitive ability and personality traits, are considered in OLS specifications, using individual fixed effects has a large impact on premium estimates. Compared to OLS estimates with all individual controls, the estimate of 7.7% for Major Urban by Fixed Effects is around 35% lower. That suggests that there are other individual differences that have not been measured through the set of control variables.

Columns (10) and (11) examine the effects of union membership and firm size on the urban wage premium. Although union membership is associated with an increase of 0.035 in the log of wages, likely due to union bargaining power, it makes no differences in the urban wage premium estimates. Adding a firm size variable, on the other hand, has large effects on urban wage premiums. Major Urban's wage premium decreases by nearly half and Other Urban's is no longer significant in Column (11). Like Brown and Medoff (1989), the study finds a positive relation between firm sizes and workers' wages as shown in Table 4. For instance, working for a 20,000 or more employee firm corresponds to a 11.4% rise in wages relative to working for a below 20-employee firm. One explanation is that larger firms employ workers with better skills, the skills that have not been explained via individual fixed effects and other skill variables in the study. However, the extent of wage differentials via firm sizes is relatively large, and likely not all the differences reflect unobserved skill differences.

Table 4 Firm size and its effects on log of wages

Firm size indicator – workers employed throughout Australia	Coefficient in Column (11) of Table 2
20 to 99	0.043[a] (0.02)
100 to 499	0.087[a] (0.02)
500 to 999	0.093[a] (0.02)
1,000 to 4,999	0.104[a] (0.02)
5,000 to 19,999	0.113[a] (0.02)
20,000 or more	0.114[a] (0.02)

Contrary to this finding, in Yankow (2006), firm size effects on wage premiums were not noteworthy. There are various possible explanations of why workers in large firms receive higher wages; for example, large firms last longer and so it is more important for them to pay high wages to retain employees. Or large firms enjoy economies of scale and are able to pay higher wages; nevertheless, that does not explain why they have to do so (Brown & Medoff, 1989). In this analysis, it is also possible that firm size effects are actually agglomeration economies because local economy sizes have not been accounted for and large firms in Australia are located in high density areas. Another possibility is that plants are larger, on average, in areas where the industry is concentrated than in other areas (Holmes & Stevens, 2002), and the relation between wages and firm size could be the benefits of localization economies. Or large firms in Australia gain more from large local economy than small firms due to better market access, and wage benefits associated with increased firm size are external to the firm (Combes & Gobillon, 2015). In short, it is not clear from this study what the positive relation between firm size and employees' wages represents, and more work is needed to address the question. Allowing for these considerations, Column (11) in Table 2 shows that movers from Rural Area to Major Urban experience a 4.6% increase in their wages relative to stayers.

The above results with specifications without individual fixed effects and Fixed Effects models show that urban wage premiums in Australia, like in other countries, do not solely result from differences in skills between urban and rural workers. Accounting for firm size effects reduce the premiums' magnitude but do not make them disappear. The wage advantage of around 20% for urban workers is relatively smaller than in other countries, but the advantage of around 7.5% after accounting for individual differences with fixed effects is relatively larger. Using Fixed Effects models, Yankow (2006) obtained a premium of around 5% for the US, and D'Costa and Overman (2014) obtained a premium of 2.5% for the UK. There are slightly less variations in skills across areas and agglomerations are potentially more important in explaining spatial wage differentials in Australia than in other countries.[20]

As studies on the topic from other countries focus on full-time males, to make the results easily compared, this study also focuses on full-time males. Even though I have not done all the analyses for females, the results of the analyses I have done (in the appendix) are in line with those for male workers. OLS estimates without individual characteristic variables show that full-time female workers in Major Urban earn 12.8% more than their counterparts in Rural Area. Fixed Effects estimates with specifications like in Column (2) in Table 2 suggest that the premiums are 6.3% and 2.8% for females in Major Urban and Other Urban

[20] As discussed in the literature review, μ_i represents innate abilities that are not captured by observed characteristics. These unobserved abilities like intelligence or family background likely relate to highest education levels achieved and job tenures. As μ_i potentially correlates with X_{it} , Fixed Effects are preferred to Random Effects: Fixed Effects estimators are consistent while Random Effects are not. In fact, Hausman tests resoundingly reject Random Effect models.

respectively. Both male and female Australian workers experience urban wage premium, and the other conclusions in this thesis are likely for both males and females.[21]

The next analysis concerns potential endogeneity bias resulting from the association between location and learning. It is the wage growth effects mentioned in Glaeser and Maré (2001). This part also examines how learning differs in urban and rural Australia.

7.2 Urban learning benefits

7.2.1 First Difference estimates

The number of migrations from Rural to Other Urban is 210 and other way around is 246. For Major Urban and Other Urban, the numbers are 188 migrations and 152 migrations. Also, between Major and Other Urban, there are more migrations from Other to Major Urban, 291 compared to 271. The urban-rural migration direction and potential urban learning benefits could lead to bias in Fixed Effects estimates (the discussion in the econometric specifications). Following Yankow (2006), I consider First Difference of individual wage equations to clear out some urban wage growth effects and check the robustness of the Fixed Effects estimates

$$\Delta y_{ict} = \Delta X_{it}\theta + \beta_{Major}\Delta Major\ Urban + \beta_{Other}\Delta Other\ Urban + \Delta\epsilon_{ict} \text{ (Equation (10))}$$

Column (1) in Table 5 presents the estimates of the urban wage premium by First Difference. Independent variables in Column (1) are the difference between years $t+1$ and t of independent variables in the specification in Column (9) in Table 2. They are changes in years of schooling, education level, years of work experience and its square, occupation and industry. The reference group includes people who do not move during the period between years t and $t+1$. The point estimate of $\Delta Major\ Urban$ is 0.033, statistically significant at the 5% level. In other words, moving from Rural Area to Major Urban is associated with an immediate wage rise of 3.3% relative to not moving, and moving in the opposite direction is associated with an immediate wage fall of 3.3%. The effect is smaller for Other Urban with an estimated wage premium of 2.5%, only statistically significant at the 10% level. The urban wage effects are weaker with the First Difference than with Fixed Effects. Compared to the respective wage premium estimates by Fixed Effects in Column (9) in Table 2, the urban wage premium estimated by First Difference is only about half. On the contrary, Yankow (2006) found that differences between Fixed Effects and First

[21] To take rural-urban differences in wages of part-time workers into the estimation, I include a 'part-time' indicator, taking the value of one for part-time work and zero otherwise, and use a sample having both part-time and full-time male workers. Other variables and the estimation method are as in Column (10) in Table 2. The results are very close to corresponding results with full-time males in Table 2: the estimated Major Urban's premium is 7.5% (p-value = 0.000) and Other Urban's premium 3.5% (p-value = 0.011).

Difference estimates of the urban wage premium are insignificant. The reduction in wage premium estimates is also not what we expected given the urban-to-rural migration direction in the data and assuming learning benefits in big cities as discussed in the econometric specifications. Even though the wage gain is smaller than expected, rural-to-urban migrants do not experience an immediate wage loss of around AUD 10 per hour associated with the move as found by Rowe et al. (2017) for young Australians.[22] Likely, career paths are more important for young workers, and they accept the loss for future career development.

[22] Different from Rowe et al. (2017) where wage increases or decreases are absolute, wage rises (increases) or falls (decreases) in the study's analyses are relative (moving from a rural area to an urban area compared to not moving or staying in rural areas, etc.).

Table 5 Estimation of urban wage premium by First Difference

	OLS (1)	OLS (2)
ΔMajor Urban	0.033[b] (0.02)	
Stay in Major Urban		–0.002 (0.00)
Move in Major Urban		0.049[b] (0.02)
Move out Major Urban		–0.019 (0.02)
Stay in Other Urban		–0.003 (0.00)
ΔOther Urban	0.025[c] (0.02)	
Stay in Other Urban		–0.002 (0)
Move in Other Urban		0.031 (0.02)
Move out Other Urban		–0.020 (0.02)
ΔMarried or de facto	0.010 (0.01)	0.009 (0.01)
ΔYears of schooling	0.019[a] (0.01)	0.019[a] (0.01)
ΔYears of experience	0.072[a] (0.02)	0.071[a] (0.02)
ΔYears of experience sq.	–0.001[a] (0.00)	–0.001[a] (0.00)
ΔPostgraduate	0.034 (0.06)	0.034 (0.06)
ΔBachelor	0.040 (0.05)	0.040 (0.05)
ΔLevel 12	0.066[a] (0.02)	0.066[a] (0.02)
ΔYears in current occupation	0.000 (0.00)	0.000 (0.00)
ΔYears with current employer	–0.001 (0.00)	–0.001 (0.00)
ΔOccupation indicator	Yes	Yes
ΔIndustry indicator	Yes	Yes
N	40,101	40,101
R sq.	0.010	0.010

All regressions include a constant term. Dependent variable is the change in the log of wages between years $t + 1$ and t, i.e. $\Delta lnw_{it} = lnw_{i(t+1)} - lnw_{it}$. Numbers in brackets are robust standard errors, clustered on individuals. [a], [b] and [c] are significant at 1%, 5% and 10% respectively.

In Column (2) in Table 5, each of $\Delta Major\ Urban$ and $\Delta Other\ Urban$ is augmented into 'Move in', and 'Move out' indicators and 'Stay in' indicators are added. Workers moving from Rural to Major Urban enjoy a 4.9% increase in wages relative to workers staying in Rural Area for the period. Other coefficients are not significantly different from zero above the 10% level even though the signs are in line with the expectation: moving into an urban area gains wages and moving out loses wages. Notably, workers moving out of Major Urban or Other Urban to Rural Area do not experience expected wage reductions. That observation is in line with Glaeser and Maré (2001)'s results with both the PSID and the NLSY datasets but not with Yankow (2006). Yankow (2006) found wage gains associated with moving in and wage losses associated with moving out of cities were equal in magnitude at around 6.5% in the US. The imbalance of wage gain and wage loss for opposite migrations (after considering other skill development) suggests that the immediate shift in wages does not totally reflect the difference between the two locations' effects. One explanation is that experience gained in Major Urban or Other Urban is well regarded in Rural Area, and so urban-rural movers do not lose wages despite losing the urban wage premium. Another explanation, of Glaeser and Maré (2001), is that workers who move out are presented with good job prospects at their destination.

The coefficients of 'Staying in Major Urban' and 'Staying in Other Urban' are not significantly different from zero, suggesting that there are no learning benefits associated with living in urban Australia. The result contradicts findings for Spain and the US where workers stay in large urban centers enjoyed higher wage growth than workers in those countries' other areas (De la Roca & Puga, 2017; Yankow, 2006). The result that there is no 'extra' urban wage growth, however, is consistent with Wheeler (2006)'s finding in which the sample was constrained to US workers who did not move during the period. If experiences in different areas have similar effects on individual capital accumulation in Australia, the good news is that the possible endogeneity due to correlation between location and local learning is of little concern in our case. Fixed Effects models will provide accurate wage premium estimates regardless of the migration direction.

The above analysis with First Difference, as with Fixed Effects, finds urban wage premiums in Australia, especially for large population centers. A more detailed analysis on migration direction suggests a rather complex relation between wage outcomes and movements. A self-selection process where migrating decisions are dependent on opportunities at the destination is possible. It is an inherent issue with using individual fixed effects and within-individual estimation and addressing it requires another approach. However, if workers move from an urban to a rural area when they have a good opportunity, this also applies to workers moving from a rural to an urban area. The first suggests a downward bias in urban wage premium estimates and the second an upward bias. Because movements are in both directions, the overall bias is probably small.[23] In the next section, I examine the transitional noise that could affect urban wage premium estimates, especially with First Difference estimates: possible Ashenfelter's Dip before migration and wage dips immediately after migration (Rowe et al., 2017).

7.2.2 Long difference models

Because of the requirements of valid observations at the beginning and the end of the period and staying in the same place after the first year, we lose observations quickly for longer lagged difference models. In the study's sample, there are 19,656 observations with valid individual wages in years t and $t+4$ and individuals stay in the same location in years $t+1$ to $t+4$. Among them, 163 migrations were to Major Urban, 155 to Other Urban, and 132 out Major Urban, 182 out Other Urban.

Table 6 Wage premium estimates by long difference models

	(1)	(2)	(3)	(4)
Dependent variable	$lnw_{t+2} - lnw_t$	$lnw_{t+3} - lnw_t$	$lnw_{t+4} - lnw_t$	$lnw_{t+2} - lnw_{t-1}$
Stay in Major Urban	–0.001	0.000	0.011	0.000
	(0.01)	(0.01)	(0.01)	(0.01)
Move in Major Urban	0.086[a]	0.118[a]	0.085[b]	0.058[c]
	(0.03)	(0.03)	(0.04)	(0.03)
Move out Major Urban	–0.038	–0.055	–0.090	–0.094[b]
	(0.03)	(0.04)	(0.06)	(0.04)
Stay in Other Urban	–0.002	–0.002	0.005	–0.001
	(0.01)	(0.01)	(0.01)	(0.01)
Move in Other Urban	0.032	0.057[c]	0.095[b]	0.124[a]
	(0.03)	(0.03)	(0.05)	(0.04)
Move out Other Urban	–0.026	–0.039	–0.029	–0.008
	(0.02)	(0.03)	(0.03)	(0.03)
N	30,829	24,230	19,170	24,189
R sq.	0.024	0.037	0.047	0.037

All regressions include a constant term. Other independent variables are in the same lagged difference as the dependent variable: married, years of schooling, years of experience and its square, education levels, tenure in current occupation and with current employer, occupation indicator and industry indicator. Numbers in brackets are robust standard errors, clustered on individuals. [a], [b] and [c] are significant at 1%, 5% and 10% respectively.

[23] This argument is like the previous argument around learning effects. More accurately, with slightly more workers moving from urban areas to rural areas in the sample, there might be a slight downward bias in urban wage premium estimates, assuming the self-selection process as mentioned.

Table 6 presents the results with long difference models. The reference group consists of stayers in Rural Area for the whole period.[24] Again, workers residing in Major Urban or Other Urban for the period – two years in Column (1), three years in Column (2) and four years in Column (3) – do not receive rises in their wages relative to workers residing Rural Area for the same period. Migrants from Rural Area to Major Urban between years t and $t+1$ experience a rise of 0.086 in the log of wages for the period from years t to $t+2$ relatively to stayers in Rural Area. The wage rises are relatively stable for longer periods in Columns (2) and (3): for the three-year period, the rise is 11.8%, and for the four-year period, the rise is 8.5% relatively to stayers in Rural Area. Movers from Rural Area to Other Urban receive a wage rise of 5.7% for the three-year period and a wage rise of 9.5% for four-year period relative to stayers in Rural Area (results in Columns (2) and (3)).

Workers moving from rural areas to Australian large urban centers (Major Urban) experience high wage growth in the year following migration (between years $t+1$ and $t+2$): the wage rise associated with the migration increases from 4.9% upon arrival to 8.6%. The results agree with the work of Rowe et al. (2017) where they found high wage growth in the years following rural-to-urban migration. It is likely that workers moving from rural areas to urban areas do not receive the urban wage premium in full upon arrival, and the wage rise of 4.9% for 'Moving in Major Urban' relative to staying in Rural Area probably underestimates the area's wage premium. Indeed, considering a year after rural-to-urban migration rather than upon arrival, the wage rise of 8.6% is close to the urban wage premium estimate of 7.5% by Fixed Effects.

It is evident from the previous results that rural-to-urban migrants experience relative wage dips upon arrival when their wages do not reflect the full urban wage premium. To examine whether the migrants also experience wage dips immediately before migration (Ashenfelter & Card, 1984; Glaeser & Maré, 2001), I compare rural-to-urban migrants' wages a year after the move to a year before the move. Column (4) in Table 6 shows that migrants from Rural Area to Major Urban experience a rise of 5.8% in their wages relatively to stayers in Rural Area. The small rise compared to 8.6% in Column (1) suggests that between years t and $t+1$, migrants experience a wage dip of around 2.8% (i.e.$\approx 8.6\% - 5.8\%$) for the year before the move (between years $t-1$ and t) relative to workers who stay in Rural for the same period. The wage dip prior to migration may also explain why movers from Rural Area to Major Urban do

[24] The reference group consists of stayers in Rural Area in years t and $t+1$ but they also stay in the same area in other years in the period following $t+1$. In effect, the reference group consists of stayers in Rural Area for the whole period.

not experience notable falls in their wages when we compare wages immediately before and after the move. If we consider a year before the move, movers from Rural Area to Major Urban experience a fall of 9.4% in their wages relative to stayers in Rural Area for the period from years $t-1$ to $t+2$. Interestingly, for Other Urban, the coefficient of 'Moving in' indicator increases to 0.124 and the 'Move out' indicator is not significantly different from zero in Column (4), suggesting that movers from Other Urban to Rural Area likely do not experience an earning dip prior to migration as in the case of movers from Major Urban to Rural Area. For reference, the results of long differences using some other periods are included in the appendix.

The above results from specifications using migration direction indicators and long difference models are suggestive of relative wage dips immediately before and after rural-to-urban migrations. Contrary to De la Roca and Puga (2017) and Yankow (2006), the acceleration of human capital in urban areas is not evidenced as the study finds no significant relations between high wage growth and urban status. It is also not clear from Table 6 that wage rises associated with rural-to-urban migrations increase over time postmigration as the learning hypothesis suggests. Therefore, the high wage growth within a year upon arrival is the realization of the urban wage premium rather than reflecting high growth of individual skills in cities.

As migrants' wages in the year of the move may not reflect individual skills and the urban wage premium, making estimating the premium via the immediate shift in wages inaccurate, in another regression, I omit observations in years t and $t+1$ if individuals change places between the two years and estimate the urban wage premium using Fixed Effects. Major Urban's wage premium is 7.1%, at the 5% significance level and Other Urban's is not significant in this case. The detailed results of the regression are included in the appendix.

The analyses using OLS, Fixed Effects and First Difference estimations find robust results of the urban wage premium in Australia. The estimation bias due to urban learnings and transitional noise is likely to be small. The premium effects are strong for large urban centers, centers from 100,000 persons. The urban wage premium is likely but lower for small urban centers. That is not surprising given that urban theories emphasize the effects of the largest population centers on individual wages (Glaeser & Maré, 2001). The above results, using the SOS structure, suggest a positive relation between the size of local economy and local wages in Australia. The next section examines this relationship in detail using Australian labor market regions (SA4), following Combes et al. (2008)'s two-step procedure.

7.3 Agglomeration economies in Australia

From 88 spatial SA4s, for this part, I drop observations in the four largest regions, the outback regions of Western Australia, Northern Territory, Queensland and South Australia. For these regions, characteristics in one location may be quite different from another within one area, making the regions' statistics not representative. I further exclude observations in another 16 areas with more than 10% of the workers employed full-time or part-time in agriculture, forestry and fishing or mining. These activities have rural characteristics, and the study focuses on urbanization. Productivity in these 'rural' areas is likely driven by natural conditions such as the quality of soil and the availability of minerals rather than the size of the local economy.[25] Excluding 'Other Territory' area (consisting of Jervis Bay, Cocos (Keeling) Islands, Christmas Island and Norfolk Island), the four largest areas and the 16 areas, there are 67 areas in the study. The restriction leaves us a sample of 47,097 valid individual-year observations.

Interestingly, the highest paid regions are not among the regions with highest employment density: workers in Perth – Inner, Sydney – North Sydney and Hornsby and Sydney – Ryde are the top earners; their average wages are above AUD 30 per hour. All regions, except regions around Australian capital cities, have local employment density below 500 persons per sq. km. More importantly, a positive correlation between an area's density and local wages is observed, as in other countries; a one percent increase in employment density is associated with around five percent increase in hourly wages.

[25] The list of the areas that have employment shares in the two industries above 10% is in the appendix.

Figure 4 Average hourly wages and employment density by the 67 SA4 areas

$\ln(hourly\ wages)_c = 0.048 \ln(employment\ density)_c + 18.471$, N = 67, R sq. = 0.407. The coefficient's robust standard errors are 0.007 and 0.034 respectively and are significant at the 1% level. Hourly wages (AUD) are the average of wages adjusted by inflation in the sample from 2001 to 2016. The employment figures used are the average for the 2001 to 2016 period (ABS, 2018d).

Table 7 presents estimates of the elasticity of wages with respect to local employment density. The main results, from Combes et al. (2008)'s two-step procedure described by Equations (11) and (12) in the econometric specifications, are presented in Columns (4) and (5) in Table 7.

$$y_{ict} = \mu_i + X_{it}\theta + \beta_c + \epsilon_{ict} \text{ (Equation (11))}$$

$$\widehat{\beta_c} = Z_c\gamma + \eta_c \text{ (Equation (12))}$$

X_{it} consists of individual characteristic variables and industry indicators used throughout the study, as shown in Table 7. In the second stage, only the log of employment density is included in Z_c, and a region's employment density is its average for the period from 2001 to 2016. Estimated elasticity of earnings with respect to employment density with the specification is 0.018, statistically significant at the 1% level, shown in Column (5). The estimated elasticity is slightly smaller but not significantly different from a value of around 0.02 typically found in other countries with similar fixed effects controls for individual heterogeneity (Combes & Gobillon, 2015). De la Roca and Puga (2017) found an elasticity of 0.024 for Spain with similar sample restrictions: full-time male workers, and similar first-stage individual controls on individual fixed effects, education and work experience. Combes et al. (2010), using fewer individual controls, namely individual fixed effects, age and age squared, obtained a higher coefficient of 0.033 for France (sea, lake and mountain are location characteristic controls in the second stage).

Column (1) in Table 7 incorporates econometric developments in cluster-robust standard errors. The log of employment density and individual characteristics are included in a one-stage OLS regression.[26] The standard errors are two-way clustered on individuals and areas rather than only on individuals. Because people change their location, individual clusters and location clusters are non-nested.[27] The relation between and local employment density and local workers' wages is statistically significant at the 1% level.

[26] The one stage regression is $y_{ict} = \gamma \ln(emp\ den)_{ct} + X_{it}\theta + \beta_c + \epsilon_{ict}$. The significance of the elasticity of individual wages with respect to the region's employment density is the main interest.

[27] Two-way clustering by Stata's *ivreg2* written by Baum, Schaffer and Stillman (2002).

Table 7 Estimations of wage elasticity with respect to employment density

	OLS (1)	OLS (2)	OLS (3)	Fix (4)	OLS (5)	Fixed (6)	OLS (7)
Dependent variable	Log of wages	Log of wages	$\widehat{\beta_c}$ from (2)	Log of wages	$\widehat{\beta_c}$ from (4)	Log of wages	$\widehat{\beta_c}$ from (6)
Log of employment density	0.028[a] (0.01)		0.024[a] (0.00)		0.018[a] (0.00)		0.018[a] (0.00)
SA4 location indicator $\widehat{\beta_c}$		Yes		Yes		Yes	
Individual fixed effects				Yes		Yes	
Married or de facto	0.110[a] (0.01)	0.090[a] (0.01)		0.044[a] (0.01)		0.044[a] (0.01)	
Years of schooling	0.004 (0.00)	0.002 (0.00)		0.020[a] (0.00)		0.020[a] (0.00)	
Years of experience	0.033[a] (0.00)	0.031[a] (0.00)		0.052[a] (0.01)		0.052[a] (0.01)	
Years of experience sq.	–0.001[a] (0.00)	–0.001[a] (0.00)		–0.001[a] (0.00)		–0.001[a] (0.00)	
Postgraduate	0.396[a] (0.04)	0.310[a] (0.03)		0.098[b] (0.05)		0.099[b] (0.05)	
Bachelor	0.312[a] (0.03)	0.220[a] (0.02)		0.056 (0.04)		0.058 (0.04)	
Level 12	0.096[a] (0.01)	0.076[a] (0.01)		0.082[a] (0.02)		0.081[a] (0.02)	
Years in current occupation	0.003[a] (0.00)	0.003[a] (0.00)		0.001[a] (0.00)		0.001[a] (0.00)	
Years with current employer	0.002[b] (0.00)	0.001[b] (0.00)		0.001 (0.00)		0.001 (0.00)	
Local industry employment share x industry indicator						Yes	
Occupation indicator	No	Yes		Yes		Yes	
Industry indicator	No	Yes		Yes		Yes	
Other time invariant controls*	Yes	Yes		N/A		N/A	
N	35,624	35,273	67	45,737	67	45,670	67
R sq.	0.374	0.458	0.402	0.393	0.231	0.394	0.234

All regressions include a constant term. Regressions in Columns (1), (2), (4) and (6) include time indicators. Numbers in brackets are robust standard errors, clustered on individuals in Columns (2), (4) and (6) and clustered (non-nested) on individual and location (SA4) in Column (1). [a], [b] and [c] are significant at 1%, 5% and 10% respectively. Reported R sq. is overall for OLS regressions and within-individual for Fixed Effects regressions.

*Other time-invariant controls in OLS regressions are born overseas, cognitive test scores (three variables: Backwards Digit Span, Symbol Digit Modalities and short NART) and personality test scores (five variables: extroversion, agreeableness, conscientiousness, emotional stability and openness).

Column (2) includes all individual characteristic variables in X_{it}, taking advantage of the HILDA's cognitive ability and personality trait variables. Individual fixed effects are not included in the first stage, and location effects for each labor market region, β_c, are estimated using OLS. The elasticity of wages with respect to density with that first-stage specification is 0.024, presented in Column (3). Thus, controlling for unobserved individual heterogeneity with fixed effects has a large impact on estimated elasticity, reducing the elasticity by 25% to 0.018 in Columns (5) and (7). The drop is greater in De la Roca and Puga (2017), at 48% from 0.046 to 0.024, but the current study includes measures of cognitive ability and personality traits in the first stage. Closer to this study's result, Combes et al. (2010) found that the elasticity fell by 35% from 0.051 to 0.033 when fixed effects were included. The finding agrees with the literature and with previous analysis using the SOS structure that unobserved individual differences explain a significant portion of wage differences between areas.

Columns (6) and (7) in Table 7 presents the results where a measure of local specialization, $emp\ share_{cst} = \frac{emp_{cst}}{emp_{ct}}$, is included in the first stage. Industry specialization effects, δ_s in Equation (13), are estimated by including the interactions between $emp\ share_{cst}$ and industry indicators. Accounting for local industry effects via specialization makes little differences in our case; estimated elasticity stays unchanged at 0.018. It appears that wage benefits from specialization are not much different among labor markets of different sizes in Australia. Small effects of specialization on the estimate of the elasticity are also found by De la Roca and Puga (2017) for Spain.[28]

The estimated relationship between workers' wages and local employment density is robust with specifications in Table 7.[29] Related literature suggests that individual skills have more important roles than location characteristics in estimates of agglomeration economies (Combes & Gobillon, 2015). To examine whether this is the case for Australia, following the literature, the study uses instruments to account for the possible endogenous local economy size. Like Combes et al. (2008, 2010) who used 'sea' as a control

[28] In another specification, I include all individual controls, local specialization, firm size and local employment density in a one-stage regression to separate the effects of firm size and density on wages. While firm size effects are due to increasing returns to scale within firms, agglomeration benefits result from clustering of economic activities in the area. The specification is similar to that of Matano and Naticchioni (2012). The effects of density on wages remain robust in this case; the elasticity of wages with respect to density is 0.019, significant at the 1% level. More details are in the appendix.

[29] The conclusion of agglomeration economies from Table 7 do not depend on the exclusion of areas that have more than 10% of the total employed in agriculture, forestry and fishing or mining. Using all 87 labor market regions and the specification in Column (4) in Table 7, the estimated elasticity is 0.036 (p-value = 0.018).

variable, I include 'distance to coast' as a control variable.[30] Areas close to seas or ports attract people due to scenery and weather and are well-connected to other areas and countries through waterways.

Table 8 Location characteristic variables statistics (67 labor market regions)

	Mean	SD	Min	Max
Estimated location effects, $\widehat{\beta_c}$	0.084	0.081	–0.177	0.317
Ln(employment density)	4.364	2.222	–0.136	7.835
Distance to coast (100 km)	0.006	0.016	0.000	0.082
Ln(population density 1911)	2.951	2.299	–1.501	8.427
Bulk density (g/cm^3)	1.365	0.064	1.163	1.492
Available water capacity (%)	13.625	0.712	11.954	15.770
Terrain ruggedness (m)	124.255	96.877	6.654	329.095

Table 9 presents the results with IV estimations. Column (1) shows that the relation between employment density and distance to coast is not significantly different from zero. On the other hand, all the instruments, namely ln(population density 1911), bulk density and available water capacity, have good explanation power over ln(employment density). The relations of the instruments with current employment density are as expected: positive for past population and negative for bulk density and terrain ruggedness. It is possible that densely grown forest in areas with high available water capacity hindered population settlements in the past, resulting in the negative relation between available water capacity and current employment density.

The underidentification LM test suggests that the instruments for ln(employment density) are relevant. In addition, those instruments are strong; the Cragg–Donald Wald F statistic is 60.445, exceeding all critical values proposed by Stock and Yogo (2005) for maximal relative IV bias and minimal IV size. The Sargan–Hansen test for overidentification restrictions cannot reject the null hypothesis that all instruments are exogenous at the 90% confidence level. Different types of instruments, i.e. historical population, soil quality and terrain characteristics, give more weight to the validity of the overidentification test because it is not likely that all the instruments are not valid.

As shown in Column (2), using instruments makes little difference to the estimate of elasticity, a finding consistent with other studies' (Combes et al., 2008, 2010; De la Roca & Puga, 2017). For example, De la Roca and Puga (2017)'s estimate of elasticity decreased by a small amount from 0.022 to 0.02 when IV

[30] The variable 'sea' in Combes et al. (2008) is the percentage of municipalities in the areas that have a sea shore.

estimation was used. In fact, the endogeneity test cannot reject the null hypothesis that current employment density is exogenous (p-value = 0.731), or IV estimation is not needed.

Table 9 IV estimates of wage elasticity

	OLS (1)	2SLS (2)
Dependent variable	Ln(employment density)	$\widehat{\beta_c}$
Instrumented ln(employment density)		0.017^{a} (0.00)
Distance to coast (100 km)	10.648 (17.27)	–0.110 (0.56)
Ln(population density 1911)	0.466^{a} (0.08)	
Bulk density (g/cm^3)	-7.033^{a} (2.37)	
Available water capacity (%)	-0.519^{a} (0.18)	
Terrain ruggedness (m)	-0.012^{a} (0.00)	
N	67	67
R sq.	0.812	0.232
P-value – Underidentification LM test (Anderson canonical correlations) (H_0: the equation is underidentified)		0.000
Weak identification test (Cragg–Donald Wald F statistic)[31] (H_0: (excluded) instruments are jointly insignificant)		60.445
P-value – Sargan–Hansen overidentification test (H_0: all instruments are exogenous)		0.112
P-value – Endogeneity test (H_0: instrumented variable is exogenous)		0.731

All regressions include a constant term. Column (1) is the first-stage regression of ln(employment density) on a set of instrument variables (excluded instruments) – the log of population density in 1911, bulk density, available water capacity and terrain ruggedness – and explanatory variable (included instrument), distance to coast. Column (2) is the second-stage regression of $\widehat{\beta_c}$ on the instrumented ln(employment density) and the explanatory variable.

Numbers in brackets are robust standard errors in Column (1) and standard errors in Column (2). [a], [b] and [c] are significant at 1%, 5% and 10% respectively.

The Cragg–Donald Wald F statistic exceeds all Stock and Yogo (2005)'s thresholds for maximal IV relative bias and minimal IV size.

[31] For the specification, Stock and Yogo (2005)'s critical values for 5% maximal IV relative bias is 16.85 and for 10% maximal IV size is 24.58.

8 Conclusions

The study finds persuasive evidence of the urban wage premium in Australia, consistent with findings around the world. The estimated wage premiums are robust, especially for Australian large urban centers, centers with populations from 100,000 persons. Workers in those centers earn around 7.5% more than workers with similar levels of skills in rural areas. Workers in small urban centers also likely receive the urban wage premium, but the premium is smaller than that of large centers. A one percent increase in local employment density results in around a 1.8% increase in local workers' wages. Another finding is that rural-to-urban migrants receive only a portion of the urban wage premium upon arrival but experience high wage growth in the year after migration. The study finds no evidence for urban learning benefits as stayers in urban areas do not experience higher wage growth than stayers in rural areas. The high wage growth in the year post rural-to-urban migration is more likely the realization of the urban wage premium rather than resulting from increases in workers' skills. Interestingly, workers moving from urban areas to rural areas do not experience a notable fall in their wages. One possible explanation is that urban-to-rural migrants are not particularly successful in urban areas, receiving low wages before deciding to move out.

Large variations in terms of area and population among labor market regions possibly affect the estimate of the elasticities. A large labor market region covers the surrounding 'empty' land in addition to its town centers. In that case, the region's density may not reflect the population density in its centers where most of the area's observations are from. As a result, there is a potential mismatch between local density and individual characteristics used in the study. Besides, the study finds that firm sizes potentially have large impact on employees' wages and explains part of the urban wage premium in Australia. It is not clear whether the positive relation between firm sizes and employees' wages results from economies of scales within firms or from external factors such as large local economies benefiting large firms more than small firms. In this study, the role of firms in the urban wage premium has not been examined appropriately, and further research is needed to address the question.

The analyses undertaken in this dissertation suggests that the widening rural-urban wage gaps observed in Australia (National Rural Health Alliance Inc., 2014) are not entirely due to differences in individual skills. An effective policy aims to minimize the gaps need consider both skills and locations. The urban wage premium is among factors that attract workers to Australian cities. We would like to know further whether the urban wage premium suggests that we should live in urban areas? To answer that question is the same as to test the spatial utility equalization assumption: do high wages compensate for high housing costs, work stress and traffic congestions in urban areas? Where to live also depends on individuals' tastes: some like socializing and competing with others while some prefer peace of mind. If utility is the same as life satisfaction, the HILDA survey provides answers for a range of life satisfaction questions that can be used to test the assumption directly. They are all very interesting topics for future research.

<u>Note</u>

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 Social Services (DSS) 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 author and should not be attributed to either DSS or the Melbourne Institute.

References

Abdel-Rahman, H., & Fujita, M. (1990). Product Variety, Marshallian Externalities, and City Sizes. *Journal of Regional Science*, *30*(2), 165–183. https://doi.org/10.1111/j.1467-9787.1990.tb00091.x

Ashenfelter, O. C., & Card, D. (1984). Using the longitudinal structure of earnings to estimate the effect of training programs. *The Review of Economics and Statistics*, *67*(4), 648–660. https://doi.org/10.2307/1924810

ASVB. (n.d.). Understanding ASVAB Scores. Retrieved June 25, 2018, from http://official-asvab.com/understand_coun.htm

Attard, B. (n.d.). The Economic History of Australia from 1788: An Introduction. Retrieved September 12, 2018, from https://eh.net/encyclopedia/the-economic-history-of-australia-from-1788-an-introduction/

Australian Bureau of Statistics. (2004). Year Book Australia, 2004, cat. no. 1301.0. Retrieved June 25, 2018, from http://www.abs.gov.au/ausstats/abs@.nsf/Previousproducts/1301.0Feature Article32004

Australian Bureau of Statistics. (2006). ANZSCO – Australian and New Zealand Standard Classification of Occupations, First Edition, Revision 1, "ANZSCO – Australian and New Zealand Standard Classification of Occupations", publication, cat. no. 1220.0. Retrieved October 13, 2018, from http://www.ausstats.abs.gov.au/ausstats/subscriber.nsf/0/69651C2DD21FE15BCA2575DF001CB1CC/$File/12200_2006.pdf

Australian Bureau of Statistics. (2010). Australian Statistical Geography Standard (ASGS): Volume 1 – Main Structure and Greater Capital City Statistical Areas, July 2011, "Statistical Area Level 4 (SA4) ASGS Ed 2011 Digital Boundaries", ESRI shapefile, cat. no. 1270.0.55.001. Retrieved July 2, 2018, from http://www.abs.gov.au/AUSSTATS/abs@.nsf/Lookup/1270.0.55.001Main+Features1July 2011?OpenDocument

Australian Bureau of Statistics. (2012a). Australian Statistical Geography Standard (ASGS): Correspondences, July 2011, "Postcode 2011 to Statistical Area Level 4 2011", Excel table, cat. no. 1270.0.55.006. Retrieved September 20, 2018, from http://www.abs.gov.au/AUSSTATS/abs@.nsf/DetailsPage/1270.0.55.006July 2011?OpenDocument

Australian Bureau of Statistics. (2012b). Australian Statistical Geography Standard (ASGS): Volume 4 – Significant Urban Areas, Urban Centres and Localities, Section of State, July 2011, cat. no. 1270.0.55.004. Retrieved July 18, 2018, from http://abs.gov.au/AUSSTATS/abs@.nsf/Previousproducts/1270.0.55.004Main Features1July 2011?opendocument&tabname=Summary&prodno=1270.0.55.004&issue=July 2011&num=&view=

Australian Bureau of Statistics. (2012c). Australian Statistical Geography Standard (ASGS): Volume 4 – Significant Urban Areas, Urban Centres and Localities, Section of State, July 2011, "Section of State (SOS) ASGS Edition 2011 Digital Boundaries", ESRI shapefile, cat. no. 1270.0.55.0. Retrieved July 20, 2018, from http://www.abs.gov.au/AUSSTATS/abs@.nsf/Lookup/1270.0.55.004Main+Features1July 2011?OpenDocument

Australian Bureau of Statistics. (2013). Population projections, Australia, 2012 to 2101, cat. no. 3222.0.

Retrieved June 25, 2018, from http://www.abs.gov.au/ausstats/abs@.nsf/Lookup/3222.0main+features52012 (base) to 2101.

Australian Bureau of Statistics. (2014a). Australian Historical Population Statistics, 2014, "Population Distribution", Excel table, cat. no. 3105.0.65.001. Retrieved September 20, 2018, from http://www.abs.gov.au/AUSSTATS/abs@.nsf/DetailsPage/3105.0.65.0012014?OpenDocument

Australian Bureau of Statistics. (2014b). Australian Population Grid 2011, ESRI grid, cat. no. 1270.0.55.007. Retrieved September 6, 2018, from http://www.abs.gov.au/AUSSTATS/abs@.nsf/DetailsPage/1270.0.55.0072011?OpenDocument

Australian Bureau of Statistics. (2016a). Australian Statistical Geography Standard (ASGS): Volume 1 – Main Structure and Greater Capital City Statistical Areas, July 2016, "Australia (AUS) ASGS Ed 2016 Digital Boundaries", ESRI shapefile, cat. no. 1270.0.55.001. Retrieved September 20, 2018, from http://www.abs.gov.au/AUSSTATS/abs@.nsf/DetailsPage/1270.0.55.001July 2016?OpenDocument

Australian Bureau of Statistics. (2016b). Australian Statistical Geography Standard (ASGS): Volume 1 – Main Structure and Greater Capital City Statistical Areas, July 2016, cat. no. 1270.0.55.001. Retrieved July 2, 2018, from http://www.abs.gov.au/AUSSTATS/abs@.nsf/Lookup/1270.0.55.001Main+Features10018July 2016?OpenDocument

Australian Bureau of Statistics. (2018a). Australian Demographic Statistics, Dec 2017, "ERP by SA2 and above (ASGS 2011), 1991 to 2016", ABS.Stat, cat. no. 3101.0. Retrieved July 20, 2018, from http://stat.data.abs.gov.au/Index.aspx?DataSetCode=ERP_QUARTERLY

Australian Bureau of Statistics. (2018b). Australian Statistical Geography Standard (ASGS). Retrieved July 16, 2018, from http://www.abs.gov.au/websitedbs/D3310114.nsf/home/Australian+Statistical+Geography+Standard+(ASGS)

Australian Bureau of Statistics. (2018c). Consumer Price Index, Australia, June 2018, 'Tables 1 and 2. CPI: All Groups, Index Numbers and Percentage Changes', Excel table, cat. no. 6401.0. Retrieved October 8, 2018, from http://www.abs.gov.au/AUSSTATS/abs@.nsf/DetailsPage/6401.0Jun 2018?OpenDocument

Australian Bureau of Statistics. (2018d). Labour Force, Australia, Detailed – Electronic Delivery, March 2018, "RM1 – Labour force status by Age, Labour market region (ASGS) and Sex, October 1998 onwards", Pivot table, cat. no. 6291.0.55.001. Retrieved August 10, 2018, from http://www.abs.gov.au/AUSSTATS/abs@.nsf/DetailsPage/6291.0.55.001March 2018

Australian Government – Department of Jobs and Small Business. (2018). Employment Region Data, "Employment by Industry Time Series", Excel table. Retrieved October 9, 2018, from http://lmip.gov.au/default.aspx?LMIP/Downloads/EmploymentRegion

Australian Government – Treasury. (2017). *Analysis of wage growth*. Retrieved from https://static.treasury.gov.au/uploads/sites/1/2017/11/p2017-t237966.pdf

Australian Qualification Framework Council. (2013). *Australian Qualifications Framework, Second Edition January 2013*. The Australian Qualifications Framework Council. Retrieved from https://www.aqf.edu.au/sites/aqf/files/aqf-2nd-edition-january-2013.pdf

Bacolod, M., Blum, B. S., & Strange, W. C. (2009). Skills in the city. *Journal of Urban Economics*, *65*(2), 136–153. https://doi.org/10.1016/j.jue.2008.09.003

Baum, C. F., Schaffer, M. E., & Stillman, S. (2002). IVREG2: Stata module for extended instrumental variables/2SLS and GMM estimation. https://doi.org/https://ideas.repec.org/c/boc/bocode/s425401.html

Berglas, E., & Pines, D. (1981). Clubs, local public goods and transportation models: A synthesis. *Journal of Public Economics*, *15*(2), 141–162. https://doi.org/10.1016/0047-2727(81)90030-X

Bill, A., Mitchell, B., & Welters, R. (2006). Job Mobility and Segmentation in Australian City Labour Markets. *CofFEE Working Paper (Centre of Full Employment and Equity, Univ. of Newcastle, Australia)*, *06-11*(06), 1–17. https://doi.org/10.1504/IJEWE.2007.019280

Bradley, R., & Gans, J. S. (1998). Growth in Australian cities. *Economic Record*, *74*(226), 266–278. https://doi.org/10.1111/j.1475-4932.1998.tb01923.x

Briant, A., Combes, P.-P., & Lafourcade, M. (2010). Dots to boxes: Do the size and shape of spatial units jeopardize economic geography estimations? *Journal of Urban Economics*, *67*(3), 287–302. https://doi.org/10.1016/j.jue.2009.09.014

Brown, C., & Medoff, J. (1989). The Employer Size-Wage Effect. *Journal of Political Economy*, *97*(5), 1027–1059.

Cai, L., & Liu, A. Y. C. (2008). Union wage effects in Australia: Is there variation along the distribution? *Economic Record*, *84*(267), 496–510. https://doi.org/10.1111/j.1475-4932.2008.00513.x

Cai, L., & Waddoups, C. J. (2011). Union Wage Effects in Australia: Evidence from Panel Data. *British Journal of Industrial Relations*, *49*(SUPPL. 2). https://doi.org/10.1111/j.1467-8543.2009.00767.x

Central Intelligence Agency. (2016). *The World Factbook 2016*. *Washington, DC: Central Intelligence Agency*. Washington. Retrieved from https://www.cia.gov/library/publications/resources/the-world-factbook/geos/gt.html

Ciccone, B. A., & Hall, R. E. (1996). Productivity and the Density of Economic Activity. *The American Economic Review*, *86*(1), 54–70.

Cobb-Clark, D. A., & Schurer, S. (2012). The stability of big-five personality traits. *Economics Letters*, *115*(1), 11–15. https://doi.org/10.1016/j.econlet.2011.11.015

Coles, M. G., & Smith, E. (1998). Marketplaces and Matching. *International Economic Review*, *39*(1), 239–254. Retrieved from https://www.jstor.org/stable/2527239

Combes, P.-P., Duranton, G., & Gobillon, L. (2008). Spatial wage disparities: Sorting matters! *Journal of Urban Economics*, *63*(2), 723–742. https://doi.org/10.1016/j.jue.2007.04.004

Combes, P.-P., Duranton, G., & Gobillon, L. (2011). The identification of agglomeration economies. *Journal of Economic Geography*, *11*(2), 253–266. https://doi.org/10.1093/jeg/lbq038

Combes, P.-P., Duranton, G., Gobillon, L., & Roux, S. (2010). Estimating agglomeration economies with history, geology, and worker effects. In E. L. Glaeser (Ed.), *Agglomeration Economics* (Vol. I, pp. 15–66). The University of Chicago Press. Retrieved from http://individual.utoronto.ca/gilles/Papers/AggloFrance.pdf

Combes, P.-P., Duranton, G., Gobillon, L., & Roux, S. (2012). Sorting and local wage and skill distributions in France. *Regional Science and Urban Economics*, *42*(6), 913–930. https://doi.org/10.1016/j.regsciurbeco.2012.11.003

Combes, P.-P., & Gobillon, L. (2015). The Empirics of Agglomeration Economies. *Handbook of Regional and Urban Economics*, *5*, 247–348. https://doi.org/10.1016/B978-0-444-59517-1.00005-2

Corcoran, J., Faggian, A., & McCann, P. (2010). Human capital in remote and rural Australia: The role of graduate migration. *Growth and Change*, *41*(2), 192–220. https://doi.org/10.1111/j.1468-2257.2010.00525.x

D'Costa, S., & Overman, H. G. (2014). The urban wage growth premium: Sorting or learning? *Regional Science and Urban Economics*, *48*, 168–179. https://doi.org/10.1016/j.regsciurbeco.2014.06.006

De la Roca, J., & Puga, D. (2017). Learning byworking in big cities. *Review of Economic Studies*, *84*(1), 106–142. https://doi.org/10.1093/restud/rdw031

Di Addario, S., & Patacchini, E. (2008). Wages and the City. Evidence from Italy. *Labour Economics*, *15*(5), 1040–1061. https://doi.org/10.1016/j.labeco.2007.09.003

Dobbie, M., MacMillan, C., & Watson, I. (2014). The returns to general experience, job and occupational tenure: A study using Australian panel data. *Applied Economics*, *46*(18), 2096–2107. https://doi.org/10.1080/00036846.2014.894632

Duranton, G., & Monastiriotis, V. (2002). Mind the Gaps: The Evolution of Regional Earnings Inequalities in the U.K., 1982-1997. *Journal of Regional Science*, *42*(2), 219–256. https://doi.org/10.1111/1467-9787.00257

Duranton, G., & Puga, D. (2004). Micro-Foundations of Urban Agglomeration Economies. *Handbook of Regional and Urban Economics*, *4*, 2063–2117. https://doi.org/10.1016/S1574-0080(04)80005-1

Geoscience Australia. (2008). GEODATA 9 Second Digital Elevation Data Version 3 and Flow Direction Grid 2008. Retrieved September 20, 2018, from http://data.bioregionalassessments.gov.au/dataset/ebcf6ca2-513a-4ec7-9323-73508c5d7b93

Glaeser, E. L. (1999). Learning in Cities. *Journal of Urban Economics*, *46*, 254–277. https://doi.org/10.1006/juec.1998.2121

Glaeser, E. L., & Maré, D. C. (2001). Cities and Skills. *Journal of Labor Economics*, *19*(2), 316–342. https://doi.org/10.1086/319563

Glaeser, E. L., & Resseger, M. G. (2010). The complementarity between cities and skills. *Journal of Regional Science*, *50*(1), 221–244. https://doi.org/10.1111/j.1467-9787.2009.00635.x

Gobillon, L. (2004). The estimation of cluster effects in linear panel models. *Processed, INED*. Retrieved from http://laurent.gobillon.free.fr/page_web/articles/gobillon_2004_art2s.pdf

Gould, E. D. (2007). Cities, workers, and wages: A structural analysis of the urban wage premium. *Review of Economic Studies*, *74*(2), 477–506. https://doi.org/10.1111/j.1467-937X.2007.00428.x

Haig, B. D. (1982). Sex Discrimination in the Reward for Skills and Experience in the Australian Labour Force. *Economic Record*, *58*(1), 1–10. https://doi.org/10.1111/j.1475-4932.1982.tb00344.x

Holmes, T. J., & Stevens, J. J. (2002). Geographic concentration and establishment scale. *Review of Economics and Statistics*, *84*(4), 682–690. https://doi.org/10.1162/003465302760556495

Jones, F. L. (1983). Sources of Gender Inequality in Income : What the Australian Census Says. *Social Forces*, *62*(1), 134–152.

Kidd, M. P., & Shanon, M. (1996). The Gender Wage Gap : A Comparison of Australia and Canada. *The Canadian Journal of Economics*, *29*(Part 1), 121–125.

Leigh, A., & Ryan, C. (2005). Estimating Returns to Education: Three Natural Experiment Techniques Compared. *Centre for Economic Policy Research, Research School of Social Sciences, Australian National University, Discussion Papers: 493, 2005, 27 Pp.* Retrieved from http://search.proquest.com/docview/56706037?accountid=17248

Marshall, A. (1890). *Principles of Economics*. *Macmillan and Co., Limited* (Vol. 1). https://doi.org/10.1093/library/s5-XVII.3.238

Matano, A., & Naticchioni, P. (2012). Wage distribution and the spatial sorting of workers. *Journal of Economic Geography*, *12*(2), 379–408. https://doi.org/10.1093/jeg/lbr013

Miller, P., & Rummery, S. (1991). Male-Female Wage Differentials in Australia: a Reassessment. *Australian Economic Papers*, *30*(56), 50–69. https://doi.org/10.1111/j.1467-8454.1991.tb00530.x

Mincer, J. A. (1974). Schooling and Earnings. In *Schooling, Experience, and Earnings* (pp. 41–63). New York: National Bureau of Economic Research, Inc.

Moulton, B. R. (1990). An Illustration of a Pitfall in Estimating the Effects of Aggregate Variables on Micro Units, *72*(2), 334–338.

Mueller, G., & Erik Plug. (2006). Estimating the Effect of Personality on Male and Female Earnings. *Industrial and Labor Relations Review*, *60*(1).

Nahm, D., Dobbie, M., & MacMillan, C. (2017). Union wage effects in Australia: an endogenous switching approach. *Applied Economics*, *49*(39), 3927–3942. https://doi.org/10.1080/00036846.2016.1273492

National Rural Health Alliance Inc. (2014). *Income inequality experienced by the people of rural and remote Australia – Submission to the Senate Inquiry into the Extent of Income Inequality in Australia*.

Peng Yu. (2004). *Return to Education for Australian Male Workers : An estimate with HILDA*. *Paper for Case Studies in Applied Econometrics*. Australian National University. Retrieved from https://melbourneinstitute.unimelb.edu.au/assets/documents/hilda-bibliography/student-essays-dissertations/PengYu.pdf

Puga, D. (2010). The magnitude and causes of agglomeration economies. *Journal of Regional Science*, *50*(1), 203–219. https://doi.org/10.1111/j.1467-9787.2009.00657.x

Rosenthal, S. S., & Strange, W. C. (2008). The attenuation of human capital spillovers. *Journal of Urban Economics*, *64*(2), 373–389. https://doi.org/10.1016/j.jue.2008.02.006

Rowe, F., Corcoran, J., & Bell, M. (2017). The returns to migration and human capital accumulation pathways: non-metropolitan youth in the school-to-work transition. *Annals of Regional Science*, *59*(3), 819–845. https://doi.org/10.1007/s00168-016-0771-8

SGS Economics and Planning. (2015). Australian Cities Accounts 2014-15. Retrieved from http://www.sgsep.com.au/application/files/3014/5542/2965/Australian_Cities_Accounts_2014_15

_FINAL.pdf

Soil and Landscape Grid of Australia. (n.d.). Soil Attribute Product Details. Retrieved September 25, 2018, from http://www.clw.csiro.au/aclep/soilandlandscapegrid/ProductDetails-SoilAttributes.html

Stock, J. H., & Yogo, M. (2005). Testing for weak instruments in linear regression. In J. H. Stock & M. Yogo (Eds.), *Identification and Inference for Econometric Models: Essays in Honor of Thomas Rothenberg* (pp. 109–120). Cambridge: Cambridge University Press. Retrieved from http://www.nber.org/papers/T0284

Strauss, E., Sherman, E. M. S., & Spreen, O. (2006). *A Compendium of Neurospsychological Tests: Administration, Norms, and Commentary* (3rd ed.). Oxford University Press.

Summerfield, M., Freidin, S., Hahn, M., Li, N., Macalalad, N., Mundy, L., … Wooden, M. (2017). HILDA User Manual -- Release 16.

Tucker, G. (1843). *Progress of the United States in Population and Wealth*. New York: Press of Hunt's Merchants' Magazine.

United Nations. (2014). *World Urbanization Prospects – The 2014 Revision*. Retrieved from http://esa.un.org/unpd/wup/Highlights/WUP2014-Highlights.pdf

Viscarra Rossel, R., Chen, C., Grundy, M., Searle, R., & Clifford, D. (2014). Soil and Landscape Grid Australia-Wide 3D Soil Property Maps (3" resolution) - Release 1. v3. CSIRO. Data Collection. https://doi.org/https://doi.org/10.4225/08/5aaf553b63215

Waddoups, C. J. (2005). Trade union decline and union wage effects in Australia. *Industrial Relations*, *44*(4), 607–624. https://doi.org/10.1111/j.1468-232X.2005.00404.x

Weber, A. F. (1899). *The Growth of Cities in the Nineteenth Century – A Study in Statistics*. *A Study in Statistics*. New York: Columbia University.

Wheeler, C. H. (2006). Cities and the growth of wages among young workers: Evidence from the NLSY. *Journal of Urban Economics*, *60*(2), 162–184. https://doi.org/10.1016/j.jue.2006.02.004

Wooden, M. (2013). *The measurement of cognitive ability in wave 12 of the HILDA Survey*.

Yankow, J. J. (2006). Why do cities pay more? An empirical examination of some competing theories of the urban wage premium. *Journal of Urban Economics*, *60*(2), 139–161. https://doi.org/10.1016/j.jue.2006.03.004

Appendices

Appendix

Years of schooling

The below notional time is from *Australian Qualifications Framework, Second Edition January 2013* of the Australian Qualifications Framework Council (AQF, 2013)

Table 10 Duration in full-time years by ASCED qualifications (AQF, 2013)

AQF/ASCED code and qualification type	AQF typical duration (full-time years)
524 Certificate Level 1	0.5–1
521 Certificate Level 2	0.5–1
514 Certificate Level 3	1–2
511 Certificate Level 4	0.5–2
421 Diploma	1–2
411 Advanced Diploma	1.5–2
413 Associate Degree	2
312 Bachelor (Pass) Degree	3–4
311 Bachelor Honours Degree	4–5 (from the commencement of a Bachelor's Degree)
221 Graduate Certificate	0.5–1
211 Graduate Diploma	1–2
120 Masters Degree	1–2
110 Doctoral Degree	3–4

The study uses the average of the two bounds. For example, the time to complete a bachelor's degree is 3.5 years. If the person reports a more general qualification type such as a Postgraduate Degree, which could be either a master's or doctoral degree, the course duration will be the average of the two, 2.5 years. I set 0.5 year duration for 'unknown – not enough information' qualification type reported in the HILDA data.

Long difference models

Table 11 presents the results of long difference models: changes in individual wages and individual characteristics are for the period between years $t-1$ and $t+1$ in Column (1), between years $t-1$ and $t+2$ in Column (2) and between years $t-1$ and $t+3$ in Column (3). The possible location change is between years t and $t+1$.

Table 11 Long difference estimation with some other periods

	(1)	(2)	(3)
Dependent variable	$lnw_{t+1} - lnw_{t-1}$	$lnw_{t+2} - lnw_{t-1}$	$lnw_{t+3} - lnw_{t-1}$
Stay in Major Urban	–0.002	0.000	0.010
	(0.01)	(0.01)	(0.01)
Move in Major Urban	0.042	0.058[c]	0.127[a]
	(0.03)	(0.03)	(0.04)
Move out Major Urban	–0.068[b]	–0.094[b]	–0.086[c]
	(0.03)	(0.04)	(0.05)
Stay in Other Urban	–0.003	–0.001	0.004
	(0.01)	(0.01)	(0.01)
Move in Other Urban	0.080[a]	0.124[a]	0.154[a]
	(0.03)	(0.04)	(0.04)
Move out Other Urban	–0.037	–0.008	–0.039
	(0.03)	(0.03)	(0.04)
N	30,823	24,189	19,138
R sq.	0.023	0.037	0.048

All regressions include a constant term. Other independent variables are in the same lagged difference as the dependent variable: married, years of schooling, years of experience and its square, education levels, tenure in current occupation and with current employer, occupation indicator and industry indicator. Numbers in brackets are robust standard errors, clustered on individuals. [a], [b] and [c] are significant at 1%, 5% and 10% respectively.

Results with observations in the year of migration excluded

Table 12 presents estimates of Fixed Effects where the sample excludes observations in year t and $t+1$ are excluded if the individual change his locations between the years.

Table 12 Fixed Effects estimation using the sample excluding the year of migration

	Fixed Effects
Major Urban	0.071[b]
	(0.03)
Other Urban	0.023
	(0.03)
Married or de facto	0.039[a]
	(0.01)
Years of schooling	0.020[a]
	(0.01)
Years of experience	0.045[a]
	(0.01)
Years of experience sq.	–0.001[a]
	(0.00)
Postgraduate	0.084
	(0.06)
Bachelor	0.033
	(0.05)
Level 12	0.070[b]
	(0.03)
Years in current occupation	0.000
	(0.00)
Years with current employer	0.001
	(0.00)
Occupation indicator	Yes
Industry indicator	Yes
N	30,298
R sq.	0.401

The regression includes a constant term and time indicators. Dependent variable is log of nominal hourly wages. Numbers in brackets are robust standard errors, clustered on individuals. [a], [b] and [c] are significant at 1%, 5% and 10% respectively. Reported R sq. is within-individual.

<u>Excluded labor market regions</u>

Table 13 shows the ASGS 2011 SA4 areas excluded from the study. Shares of agriculture, forestry and fishing or mining is the total employed (full-time and part-time) in the industry in the area divided by the total employed in the area. The shares reported are the average of the shares in May 2013, May 2017 and May 2018 (averages of the preceding four quarters). The data is from 'Employment by Industry Time Series' from the Australian Labor Market Information Portal (Australian Government – Department of Jobs and Small Business, 2018, based on ABS, 2018d).

Table 13 Excluded ASGS 2011 SA4

ASGS 2011 SA4	Industry	Industry share
Western Australia – Wheat Belt	Agriculture, Forestry and Fishing	0.255
Warrnambool and South West	Agriculture, Forestry and Fishing	0.231
Western Australia – Outback	Mining	0.205
Far West and Orana	Agriculture, Forestry and Fishing	0.197
Shepparton	Agriculture, Forestry and Fishing	0.181
North West	Agriculture, Forestry and Fishing	0.167
Barossa – Yorke – Mid North	Agriculture, Forestry and Fishing	0.161
South Australia – Outback	Agriculture, Forestry and Fishing	0.159
South East	Agriculture, Forestry and Fishing	0.157
South Australia – South East	Agriculture, Forestry and Fishing	0.152
Mackay	Mining	0.149
Queensland – Outback	Mining	0.147
Darling Downs – Maranoa	Agriculture, Forestry and Fishing	0.138
Latrobe – Gippsland	Agriculture, Forestry and Fishing	0.122
Mandurah	Mining	0.121
New England and North West	Agriculture, Forestry and Fishing	0.118
Riverina	Agriculture, Forestry and Fishing	0.115
Murray	Agriculture, Forestry and Fishing	0.109
Hunter Valley exc Newcastle	Mining	0.108
Queensland – Outback	Agriculture, Forestry and Fishing	0.099

Estimated location effects $\widehat{\beta_c}$

Table 14 presents $\widehat{\beta_c}$ from specifications in Table 7. Column (1) presents $\widehat{\beta_c}$ by OLS in Column (2) in Table 7. Column (2) presents $\widehat{\beta_c}$ by Fixed Effects in Column (4) in Table 7. Column (3) presents $\widehat{\beta_c}$ when local specializations are included; the specification is in Column (6) in Table 7.

Table 14 Estimated location effects

ASGS 2011 SA4		$\widehat{\beta_c}$		
		OLS (1)	Fixed (2)	Fixed (3)
101	Capital Region	0.000	0.000	0.000
102	Central Coast	0.172	0.194	0.190
103	Central West	0.057	–0.018	–0.009
104	Coffs Harbour – Grafton	0.023	–0.075	–0.064
107	Illawarra	0.151	0.084	0.087
108	Mid North Coast	–0.035	0.073	0.073
111	Newcastle and Lake Macquarie	0.234	0.317	0.319
112	Richmond – Tweed	0.029	–0.038	–0.034
114	Southern Highlands and Shoalhaven	0.062	0.066	0.067

115	Sydney – Baulkham Hills and Hawkesbury	0.215	0.000	–0.001
116	Sydney – Blacktown	0.175	0.095	0.096
117	Sydney – City and Inner South	0.234	0.157	0.163
118	Sydney – Eastern Suburbs	0.311	0.230	0.236
119	Sydney – Inner South West	0.128	0.180	0.182
120	Sydney – Inner West	0.231	0.123	0.127
121	Sydney – North Sydney and Hornsby	0.332	0.119	0.123
122	Sydney – Northern Beaches	0.191	0.113	0.115
123	Sydney – Outer South West	0.127	0.162	0.163
124	Sydney – Outer West and Blue Mountains	0.123	0.130	0.128
125	Sydney – Parramatta	0.157	0.138	0.139
126	Sydney – Ryde	0.354	0.066	0.072
127	Sydney – South West	0.144	0.115	0.113
128	Sydney – Sutherland	0.222	0.128	0.130
201	Ballarat	0.009	–0.126	–0.126
202	Bendigo	0.027	0.010	0.009
203	Geelong	0.102	0.056	0.053
204	Hume	0.103	0.100	0.101
206	Melbourne – Inner	0.243	0.119	0.123
207	Melbourne – Inner East	0.149	0.108	0.111
208	Melbourne – Inner South	0.267	0.126	0.127
209	Melbourne – North East	0.185	0.075	0.076
210	Melbourne – North West	0.147	0.118	0.118
211	Melbourne – Outer East	0.152	0.091	0.092
212	Melbourne – South East	0.192	0.143	0.143
213	Melbourne – West	0.170	0.092	0.095
214	Mornington Peninsula	0.088	0.092	0.091
301	Brisbane – East	0.174	0.127	0.128
302	Brisbane – North	0.200	0.077	0.076
303	Brisbane – South	0.185	0.069	0.074
304	Brisbane – West	0.228	0.046	0.047
305	Brisbane Inner City	0.253	0.095	0.097
306	Cairns	0.140	0.112	0.116
308	Fitzroy	0.277	0.212	0.207
309	Gold Coast	0.129	–0.004	–0.005
310	Ipswich	0.057	0.028	0.030
311	Logan – Beaudesert	0.204	0.100	0.100
313	Moreton Bay – North	0.139	0.074	0.071
314	Moreton Bay – South	0.140	0.090	0.089
316	Sunshine Coast	0.124	0.081	0.078
317	Toowoomba	0.146	0.012	0.012
318	Townsville	0.183	0.086	0.088
319	Wide Bay	0.077	–0.094	–0.086
401	Adelaide – Central and Hills	0.088	0.098	0.100

402	Adelaide – North	0.089	0.115	0.117
403	Adelaide – South	0.121	0.122	0.123
404	Adelaide – West	0.167	0.100	0.104
501	Bunbury	0.149	0.106	0.105
503	Perth – Inner	0.368	0.131	0.130
504	Perth – North East	0.193	0.043	0.035
505	Perth – North West	0.223	0.110	0.106
506	Perth – South East	0.171	0.034	0.031
507	Perth – South West	0.258	0.135	0.134
601	Hobart	0.104	0.032	0.033
602	Launceston and North East	0.037	–0.177	–0.175
604	West and North West	0.077	–0.024	–0.021
701	Darwin	0.178	0.235	0.234
801	Australian Capital Territory	0.266	0.112	0.103

Estimation for full-time female workers

	OLS (1)	OLS (2)	Fixed Effects (3)	Fixed Effects (4)	First Difference (5)
Major Urban	0.128[a]	0.032[c]	0.063[a]	0.055[a]	0.021
	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)
Other Urban	0.005	–0.015	0.028[c]	0.029[c]	0.006
	(0.02)	(0.02)	(0.01)	(0.02)	(0.02)
Married or de facto		0.035[a]	0.041[a]	0.035[a]	0.017[b]
		(0.01)	(0.01)	(0.01)	(0.01)
Born overseas		0.018			
		(0.02)			
Years of schooling		0.003	0.011[a]	0.009[c]	0.020[a]
		(0.00)	(0.00)	(0.00)	(0.01)
Years of experience		0.022[a]	0.051[a]	0.047[a]	0.082[a]
		(0.00)	(0.01)	(0.01)	(0.02)
Years of experience sq.		0.000[a]	–0.001[a]	–0.001[a]	–0.001[a]
		(0.00)	(0.00)	(0.00)	(0.00)
Postgraduate		0.253[a]	0.097[b]	0.071	0.092
		(0.03)	(0.04)	(0.05)	(0.07)
Bachelor		0.183[a]	0.061[c]	0.018	0.059
		(0.03)	(0.04)	(0.04)	(0.07)
Level 12		0.057[a]	0.026	–0.023	0.040
		(0.02)	(0.02)	(0.03)	(0.03)
Years in current occupation		0.003[a]	0.001[b]	0.000	0.000
		(0.00)	(0.00)	(0.00)	(0.00)
Years with current employer		0.003[a]	0.001	0.000	–0.001
		(0.00)	(0.00)	(0.00)	(0.00)
Backwards Digit Span		–0.001			
		(0.00)			
Symbol Digit Modalities		0.002[a]			
		(0.00)			
Sort NART		0.007[a]			
		(0.00)			
Extroversion		0.007			
		(0.00)			
Agreeableness		–0.023[a]			
		(0.01)			
Conscientiousness		0.021[a]			
		(0.01)			
Emotional stability		0.008			
		(0.01)			
Openness to experience		–0.006			
		(0.01)			
Union		0.013	0.013c	0.015c	
		(0.01)	0.01	0.01	
Firm size indicator	No	Yes	No	Yes	No
Occupation indicator	No	Yes	Yes	Yes	Yes
Industry indicator	No	Yes	Yes	Yes	Yes
N	33,136	17,060	32,275	21,720	22,187
R sq.	0.179	0.467	0.38	0.39	0.008

Regressions (1), (2) and (5) include a constant term; regressions (3) and (4) include time indicators. Dependent variable is log of nominal hourly wages. Numbers in brackets are robust standard errors, clustered on individuals. [a], [b] and [c] are significant at 1%, 5% and 10% respectively. Reported R sq. is overall for OLS and First Difference regressions and within-individual for Fixed Effects regressions.

Table 15 Estimation of wage premiums for full-time female workers

OLS estimation with firm size

Table 16 OLS estimation with firm size

	OLS
Dependent variable	Log of wages
Log of employment density	0.019[a]
	(0.00)
Married or de facto	0.074[a]
	(0.01)
Years of schooling	0.001
	(0.00)
Years of experience	0.029[a]
	(0.00)
Years of experience sq.	0.000[a]
	(0.00)
Postgraduate	0.315[a]
	(0.03)
Bachelor	0.229[a]
	(0.03)
Level 12	0.073[a]
	(0.02)
Years in current occupation	0.004[a]
	(0.00)
Years with current employer	0.001
	(0.00)
Local industry employment share x industry indicator	Yes
Firm size indicator	Yes
Occupation indicator	Yes
Industry indicator	Yes
Other time invariant controls*	Yes
N	22,908
R sq.	0.500

The regression includes a constant term and time indicators. Dependent variable is the log of nominal hourly wages. Numbers in brackets are robust standard errors, clustered on individuals. [a], [b] and [c] are significant at 1%, 5% and 10% respectively.

*Other time-invariant controls in OLS regressions are born overseas, cognitive test scores (three variables: Backwards Digit Span, Symbol Digit Modalities and short NART) and personality test scores (five variables: extroversion, agreeableness, conscientiousness, emotional stability and openness).

Data appendix

Table 17 Bulk density and available water capacity descriptions (Soil and Landscape Grid of Australia, n.d.)

Soil property	Attribute description	Units
Bulk density – whole earth	*Bulk density of the whole soil (including coarse fragments) in mass per unit volume by a method equivalent to the core method*	g/cm^3
Available water capacity	*Available water capacity computed for each of the specified depth increments*	%

Table 18 Location characteristics by labor market regions

	ASGS 2011 SA4	Ln(employment density)	Distance to coast (100km)	Ln(population density in 1911)	Bulk density (g/cm^3)	Available water capacity (%)	Terrain ruggedness (m)
101	Capital Region	0.669	0.000	0.702	1.421	13.490	329.095
102	Central Coast	4.405	0.000	1.697	1.322	13.089	97.422
103	Central West	0.278	462.682	0.919	1.491	13.239	264.036
104	Coffs Harbour – Grafton	1.407	0.000	1.220	1.295	13.973	301.490
107	Illawarra	4.386	0.000	3.010	1.300	13.505	174.319
108	Mid North Coast	1.374	0.000	1.027	1.254	13.645	266.069
111	Newcastle and Lake Macquarie	5.219	0.000	4.561	1.336	13.474	70.958
112	Richmond – Tweed	2.247	0.000	1.803	1.277	13.989	170.333
114	Southern Highlands and Shoalhaven	2.118	0.000	1.538	1.354	13.837	262.092
115	Sydney – Baulkham Hills and Hawkesbury	3.570	0.000	1.393	1.353	13.271	157.914
116	Sydney – Blacktown	6.365	91.486	3.824	1.393	14.157	18.118
117	Sydney – City and Inner South	7.835	0.000	8.427	1.302	13.394	13.154
118	Sydney – Eastern Suburbs	7.826	0.000	7.264	1.310	13.241	20.158
119	Sydney – Inner South West	7.306	0.000	6.210	1.381	13.859	16.535
120	Sydney – Inner West	7.746	0.000	7.945	1.374	13.963	7.699
121	Sydney – North Sydney and Hornsby	6.642	0.000	5.585	1.341	13.196	51.757
122	Sydney – Northern Beaches	6.287	0.000	4.495	1.322	13.145	51.542
123	Sydney – Outer South West	4.512	54.190	2.251	1.437	13.335	119.482
124	Sydney – Outer West and Blue Mountains	3.628	217.964	2.532	1.311	13.809	304.796
125	Sydney – Parramatta	7.022	0.000	5.564	1.398	14.004	23.685
126	Sydney – Ryde	7.134	0.000	5.090	1.388	13.831	42.401
127	Sydney – South West	5.672	17.863	3.244	1.402	14.050	30.121
128	Sydney – Sutherland	5.992	0.000	2.281	1.354	13.178	59.325
201	Ballarat	1.852	202.536	2.709	1.415	14.269	134.472
202	Bendigo	1.717	570.279	2.344	1.474	12.810	122.829
203	Geelong	3.259	0.000	2.653	1.378	13.929	99.194
204	Hume	0.806	545.661	1.200	1.317	13.286	327.930
206	Melbourne – Inner	7.676	0.000	8.072	1.367	14.184	18.400
207	Melbourne – Inner East	7.109	55.859	5.846	1.412	13.802	27.958
208	Melbourne – Inner South	7.116	0.000	5.982	1.336	13.392	15.874
209	Melbourne – North East	4.761	103.554	2.895	1.347	13.531	190.508
210	Melbourne – North West	4.493	78.911	3.465	1.379	14.243	176.162
211	Melbourne – Outer East	4.918	137.619	2.502	1.271	13.646	244.548
212	Melbourne – South East	5.087	0.000	2.595	1.352	13.830	89.802
213	Melbourne – West	5.306	0.000	3.560	1.414	13.511	83.911
214	Mornington Peninsula	5.016	0.000	2.583	1.343	14.120	49.637
301	Brisbane – East	4.850	0.000	2.008	1.306	13.279	41.918
302	Brisbane – North	6.288	0.000	4.371	1.366	13.520	18.195
303	Brisbane – South	6.447	0.000	4.946	1.380	13.670	23.423
304	Brisbane – West	5.833	0.000	4.008	1.360	13.183	88.341

305	Brisbane Inner City	7.442	0.000	7.381	1.380	13.623	21.489
306	Cairns	1.612	0.000	0.758	1.354	13.645	297.608
308	Fitzroy	–0.136	0.000	–0.839	1.424	14.845	156.534
309	Gold Coast	4.901	0.000	1.881	1.240	13.216	236.361
310	Ipswich	2.935	0.000	2.249	1.372	14.263	147.732
311	Logan – Beaudesert	3.949	3.853	1.202	1.358	14.096	147.637
313	Moreton Bay – North	3.084	0.000	1.621	1.350	13.724	174.860
314	Moreton Bay – South	4.689	0.000	1.224	1.325	13.215	146.270
316	Sunshine Coast	3.823	0.000	0.861	1.289	13.099	142.727
317	Toowoomba	3.357	533.797	2.517	1.313	15.770	177.376
318	Townsville	0.265	0.000	–0.461	1.492	14.704	186.677
319	Wide Bay	0.771	0.000	0.680	1.420	14.014	154.978
401	Adelaide – Central and Hills	4.570	51.245	4.518	1.470	13.235	109.120
402	Adelaide – North	5.252	0.000	2.955	1.456	12.896	86.616
403	Adelaide – South	5.538	0.000	3.745	1.446	12.902	105.944
404	Adelaide – West	6.467	0.000	6.032	1.445	12.892	6.654
501	Bunbury	1.152	0.000	0.027	1.414	12.883	91.885
503	Perth – Inner	6.852	0.000	6.639	1.412	12.007	12.663
504	Perth – North East	4.163	0.000	1.984	1.400	12.844	104.640
505	Perth – North West	5.667	0.000	–1.501	1.392	12.336	20.895
506	Perth – South East	4.716	0.000	1.529	1.406	12.276	124.567
507	Perth – South West	5.653	0.000	3.600	1.401	11.954	12.557
601	Hobart	4.060	0.000	3.576	1.277	14.520	197.907
602	Launceston and North East	1.170	0.000	1.386	1.252	15.543	293.933
604	West and North West	0.755	0.000	0.829	1.163	14.926	247.035
701	Darwin	3.091	0.000	–1.236	1.463	14.836	14.537
801	Australian Capital Territory	4.430	821.812	0.236	1.383	13.748	300.257