THE NEGATIVE EFFECTS OF URBAN AGGLOMERATION ON HOUSING AFFORDABILITY IN AUSTRALIA

Akshay Vij*
Institute for Choice, University of South Australia, Adelaide, SA 5000, Australia. Email: vij.akshay@gmail.com

Jeffery D. Connor
School of Commerce, University of South Australia, Adelaide, SA 5000, Australia. Email: Jeff.Connor@unisa.edu.au

Andrew Beer
Business School, University of South Australia, Adelaide, SA 5000, Australia. Email: Andrew.Beer@unisa.edu.au

*Corresponding Author

Abstract: This study assesses the impacts of continued population growth in the largest Australian cities on housing affordability. Using data from the Australian census and other complementary sources over the period 2001-16, we estimate a system of seemingly unrelated and spatially lagged regressions to identify the relationship between a city's population size on the one hand, and average wages and housing costs on the other, while controlling for the confounding influence of other geographic, demographic and economic determinants. We find that annual home sales values have risen roughly thirty times faster with population than annual full-time wages across Australia. An increase in the population of an urban area by 100,000 would increase annual full-time wages by roughly $150 and annual home sales values by roughly $4,800. Our analysis also finds that real wages have not kept up with the high costs of living in large cities. For example, our model predicts that, ceteris paribus, price-to-income ratios (PIRs) in Greater Newcastle could rise from 7.0 to 8.4, if the city grows to the size of Sydney. And PIRs in Sydney themselves could rise from 13.6 to 14.8 by 2056, if the city grows to its expected size of 9.2 million. Relatedly, we find that there are no wage benefits to urban areas situated in close proximity to a large metropolitan centre, but these areas are more likely to have expensive local housing markets due to spatial spillover effects.

KEYWORDS: Agglomeration effects; housing affordability; regional growth

ACKNOWLEDGEMENTS: The authors would like to thank the Regional Australia Institute for providing funding support for this study, and Ben

Vonthethoff, Kevin Le and Jack Archer from the institute for their useful feedback on different aspects of this study. The authors would also like to thank the two anonymous referees for their helpful comments and suggestions.

1. INTRODUCTION

It is commonly accepted that bigger cities are more economically efficient than smaller cities (Glaeser 2011; Fujita *et al.*, 2001). Economies of agglomeration imply that as more firms in related industries cluster together, their costs of production decline significantly, because of improved access to labour and supporting technologies and services, knowledge spillovers between firms, and the creation of local markets (Florida *et al.*, 2017; Duranton and Puga, 2004; Quigley, 1998; Jacobs, 1969; Marshall, 1890). This can lead to improved economic outcomes in terms of average income, employment growth, innovation (Glaeser 2011) and productivity levels (Rosenthal and Strange, 2004). However, increased agglomeration might also lead to unaffordable housing, traffic congestion, and other negative externalities that cause diseconomies to set in (Richardson, 1995). The tension between agglomeration economies and urban costs serves as a useful basis for understanding the growth and decline of urban regions (Fujita, 1989; Henderson, 1974; Mills, 1967).

National and state governments in Australia have frequently adopted the view that significant positive agglomeration economies will see the biggest cities grow even bigger, and any urban costs can potentially be offset through improved planning and management of housing markets, infrastructure systems, and other supporting services. As a result, to date, Australia's long-term population growth strategy has prioritized major state capitals as the centres of greatest growth (McGurik and Argent, 2011). Between 2016 and 2056, the Australian population is projected to increase by up to 18.3 million people (Australian Bureau of Statistics [ABS], 2008), and approximately 56 per cent of this growth is expected to occur in Australia's two largest cities – Sydney and Melbourne.

The largest Australian cities are already under great strain, as reflected by expensive local housing markets and rising levels of housing unaffordability. Nationwide dwelling price-to-income ratios (PIRs) have nearly doubled between 1980 and 2018 (Organisation for Economic Co-operation and Development [OECD], 2019). Average housing costs in Sydney and Melbourne are between two and three times higher than those in other Australian cities and regional centres (Chappell and Campbell, 2018). The limited availability of affordable housing has been attributed to a variety of supply and demand-side factors that include housing tax policy, land-use planning and zoning laws, foreign investment in real estate, declining government support for social housing, changes in residential preferences, etc. (Lawson *et al.,* 2019; Eccleston *et al.,* 2018; Gurran *et al.,* 2018).

This study examines the impacts of continued population growth in the largest Australian cities on housing affordability. We estimate a system of seemingly unrelated and spatially lagged regressions to identify the relationship between a city's population size on the one hand, and average wages and housing costs on the other, while controlling for the confounding influence of other geographic, demographic and economic determinants. Data for our analysis comes from the Australian census and other complementary sources. It spans the period 2001-16, and it covers 384 consolidated urban areas across the country.

Our analysis finds that annual home sale values have risen roughly thirty times faster with population than annual full-time wages across Australia. An increase in the population of an urban area by 100,000 would increase annual full-time wages by roughly $150 and annual home sales values by roughly $4,800. Our analysis also finds that real wages have not kept up with the high costs of living in large cities. For example, our model predicts that, *ceteris paribus*, PIRs in Greater Newcastle, which currently has a population of roughly 0.5 million, could rise from 7.0 to 8.4, if the city grows to the size of Sydney, a population of 4.5 million. And PIRs in Sydney themselves could rise from 13.6 to 14.8 by 2056, if the city grows to its expected size of 9.2 million. Other factors, such as taxation policy reform in 2001, have contributed to escalating housing costs over the past two decades, but there is clear evidence that wages have failed to keep pace with house prices (Daniel *et al.,* 2018).

The study contributes to the existing literature in three ways. First, our findings contribute to the current debate on housing policy in Australia. As mentioned previously, earlier studies that have examined the scarcity of affordable housing in Australia have focused on factors other than population size, such as housing tax policy, land-use planning and zoning laws, foreign investment in real estate, declining government support for social housing, changes in residential preferences, seasonality effects, etc. (Lawson *et al.,* 2019; Eccleston *et al.,* 2018; Gurran *et al.,* 2018; Valadkhani *et al.*, 2017). The present study applies perspectives from new economic geography (Krugman, 1991; Glaeser and Mare, 2001) to offer additional insights on the extent to which the size of urban areas is a contributing factor to house price growth and unaffordable housing.

Second, our findings contribute to the literature on new economic geography. The relationship between population size, wages and housing costs has been studied previously in other contexts, with relatively similar results. For example, in their examination of urban areas in the United States, Glaeser (1998) found that a 1 per cent increase in a city's population size would increase earnings by 0.10 per cent and cost of living by 0.16 per cent. These previous examinations have worked under the assumption that housing and labour markets are perfectly competitive, and argued that residents wouldn't live in large expensive cities unless they received other external benefits, such as the ability to acquire new skills and knowledge (Glaeser *et al.,* 2016). However, there is extensive evidence to indicate that frictions in both labour and housing markets are substantial (Kaufman, 2007; Anenberg, 2016), and that young adults are increasingly unwilling to living in exceedingly large urban areas (Adamy and Overberg, 2019). In light of this evidence, the present study points to the growing gap between wages and housing costs in large urban areas as a potential explanation for their declining attractiveness, rather than evidence of their other non-pecuniary benefits.

Third, our findings offer new insights on the determinants of wage growth and housing costs, other than city size. For example, we find that urban areas with local populations that are less diverse in terms of their employment are likely to have higher housing costs. This is consistent with findings from previous studies that have examined economic gentrification and segregation effects, where expensive local housing markets force residents working in low-income industries to leave, and the urban area becomes more homogenous over time (Kohn, 2013; Lees, 2008). Similarly, we find that spatial lag is small and statistically weak in the case of wages but large, positive and statistically significant in the case of housing costs. This implies that there are no wage benefits to urban areas situated in close proximity to a large metropolitan centre, but these areas are more likely to have expensive local housing markets due to spatial spillover effects.

The remainder of this article is structured as follows: Section 2 describes the data that we used for our analysis and provides some essential background to the empirical context. Section 3 describes the econometric framework, and Section 4 presents estimation results. Section 5 concludes with a summary of key findings from our analysis.

2. DATA

Our analysis examines the empirical relationship between city size and the following two dependent variables – annual full-time wages and average home sales value. All dependent and independent variables used in our analysis are listed in Table 1. Wage data was sourced from the Australian census, and home sales data was sourced from CoreLogic. We have time series data for both measures over the following years: 2001, 2006, 2011 and 2016. All data is available at the Local Government Area (LGA) level. We used the 2016 boundaries to define each LGA in our sample, and we used appropriate correspondences provided by the ASGS to ensure that the same geographic units were being compared over time. However, multiple LGAs belonging to the same urban area were consolidated, with guidance from the significant urban areas' definition employed by the Australian Statistical Geography Standard. In all, we have 384 consolidated urban areas in our sample, and these comprise our unit of analysis throughout.

Table 2 enumerates the twenty urban areas with the highest PIRs in 2016. The list expectedly includes some of the largest Australian cities, namely Sydney, Melbourne, Gold Coast – Tweed Heads, Sunshine Coast – Noosa, Central Coast and Wollongong – Shellharbour. Perhaps surprisingly, though, the list also includes many smaller regional centres, and is notably topped by Kingston, a small coastal town in South Australia that receives a large number of tourists throughout the year. Like Kingston, most of these smaller centres are tourist towns with low wages and high housing costs, frequently in close proximity to a major metropolitan centre and places of significant 'second home' buying by metropolitan residents (Paris, 2010). Similarly, Table 3 enumerates the twenty urban areas with the lowest PIRs in 2016. Again, expectedly the list includes small towns in regional Australia, often in remote parts of the country with large proportions of indigenous people (e.g. Walgett, NSW), and/or in close proximity to former mines (e.g. Leonara, WA and Croydon, QLD). Finally, Figure 1 plots PIRs across Australian urban areas in 2016. Consistent with Tables 2 and 3, the highest PIRs are observed in the largest urban areas, and/or urban areas in close proximity to a large urban area.

Table 1 also enumerates the list of independent explanatory variables that we will be using to explain metropolitan and regional differences in wages and housing costs, along with the sources for the data. Most of these variables capture major demographic differences across different urban areas, as denoted by corresponding differences in age, education, employment and industry mix.

Table 1. Description of Variables Used in Our Analysis. Source: the Authors.

Measure	**Data description and source**
Wages	Average full-time wage data at the Local Government Area (LGA) level, collected as part of the census conducted during the following years: 2001, 2006, 2011 and 2016
Housing affordability	The average sales price of all recorded house transactions at the LGA level, collected by CoreLogic for the years: 2001, 2006, 2011 and 2016
Population size	Population data at the LGA level collected as part of the census, geographic area data at the LGA level available as part of the 'Regional Population Growth, Australia' series from the ABS
Population density	
Remoteness	Five level designation at the LGA level developed by the Australian Statistical Geography Standard (ASGS): 1 - Major cities of Australia; 2 - Inner regional Australia; 3 - Outer regional Australia; 4 - Remote Australia; and 5 - Very remote Australia.
Average age	Collected as part of the census
Labour force participation rates	
Proportion of population with a Bachelor's degree	
HHI index of industry diversity	A Herfindahl-Hirschman Index (HHI) of the distribution of employees across 19 industry types, as designated by the census, constructed at the LGA level using Census data. The lower the HHI, the more diverse the local economy.
Proportion of population that works in the following eight sectors: (1) Agriculture, forestry and fishing; (2) Mining; (3) Manufacturing; (4) Other blue collar; (5) New economy services; (6) Health care and social assistance; (7) Accommodation and food services; and (8) Other services	These variables are constructed using the 19-industry Australian and New Zealand Standard Industrial Classification (ANZSIC) system as follows: 'Other blue collar' include the following four industry categories: (1) Electricity, gas, water and waste services; (2) Construction; (3) Wholesale trade; and (4) Transport, postal and warehousing. 'New economy services' include the following three industry categories: (1) Information media and telecommunications; (2) Financial and insurance services; and (3) Professional, scientific and technical services. 'Other services' include the following eight industry categories: (1) Rental, hiring and real estate services; (2) Administrative and support services; (3) Public administration and safety; (4) Education and training; (5) Arts and recreation services; (6) Retail trade; and (7) Other services. Other industries directly correspond to their ANZSIC industry categories with the same name.

Table 2. Twenty Urban Areas in Australia in 2016 with the Highest Price-to-Income Ratios. Source: Population data has been sourced from the 2016 Census; wage data corresponds to annual wages for those that are employed full-time, and has also been sourced from the 2016 Census; and house sales value data denotes the average sales price of all recorded house transactions during 2016, and has been sourced from CoreLogic.

Urban area	**Population**	**Annual wage ($)**	**House sales value ($)**	**Price-to-income ratio**
Kingston, SA	2,349	54,132	801,145	14.8
Sydney, NSW	4,496,196	88,335	1,200,827	13.6
Byron, NSW	31,556	64,740	761,479	11.8
Melbourne, VIC	4,415,391	79,164	870,687	11.0
Queenscliffe, VIC	2,853	67,652	727,000	10.7
Northern Peninsula Area, QLD	2,796	72,072	715,986	9.9
Kiama, NSW	21,464	78,156	774,541	9.9
Wollongong – Shellharbour, NSW	272,090	63,779	604,908	9.5
Wingecarribee, NSW	47,882	71,656	662,645	9.2
Chittering, WA	5,472	68,848	576,791	8.4
Mornington, VIC	1,143	71,344	592,702	8.3
Central Coast, NSW	327,736	67,028	555,937	8.3
Bellingen, NSW	12,668	55,432	451,375	8.1
Ballina – Lismore, NSW	84,925	58,910	471,357	8.0
Surf Coast, VIC	29,397	83,096	656,812	7.9
Port Macquarie – Hastings, NSW	78,539	61,568	482,745	7.8
Gold Coast – Tweed Heads, QLD/NSW	647,092	73,346	568,181	7.7
Denmark, WA	5,845	58,604	452,750	7.7
Bega Valley, NSW	33,253	53,716	412,270	7.7
Sunshine Coast – Noosa, QLD	346,516	70,206	536,092	7.6

Table 3. Twenty Urban Areas in Australia in 2016 with the Lowest Price-to-Income Ratios. Source: Population data has been sourced from the 2016 Census; wage data corresponds to annual wages for those that are employed full-time, and has also been sourced from the 2016 Census; and house sales value data denotes the average sales price of all recorded house transactions during 2016, and has been sourced from CoreLogic.

Urban area	Population	Annual wage ($)	House sales value ($)	Price-to-income ratio
Yalgoo, WA	337	98,852	30,000	0.3
Dundas, WA	772	81,744	35,792	0.4
Barcoo, QLD	267	78,624	35,000	0.4
Menzies, WA	490	87,360	45,000	0.5
Paroo, QLD	1,640	58,708	35,125	0.6
Mount Magnet, WA	482	82,680	59,375	0.7
Cue, WA	194	91,572	68,981	0.8
Koorda, WA	414	83,668	66,166	0.8
Central Darling, NSW	1,833	62,868	51,458	0.8
Yilgarn, WA	1,202	72,124	59,500	0.8
Leonora, WA	1,411	104,260	88,333	0.8
Mukinbudin, WA	555	87,620	75,313	0.9
Croydon, QLD	294	68,588	61,042	0.9
Sandstone, WA	89	67,704	61,000	0.9
Trayning, WA	350	64,636	60,250	0.9
Meekatharra, WA	1,067	89,804	84,583	0.9
Etheridge, QLD	799	84,240	79,417	0.9
Walgett, NSW	6,107	60,060	58,060	1.0
Kulin, WA	765	68,536	66,979	1.0
Mount Marshall, WA	527	67,964	71,458	1.1

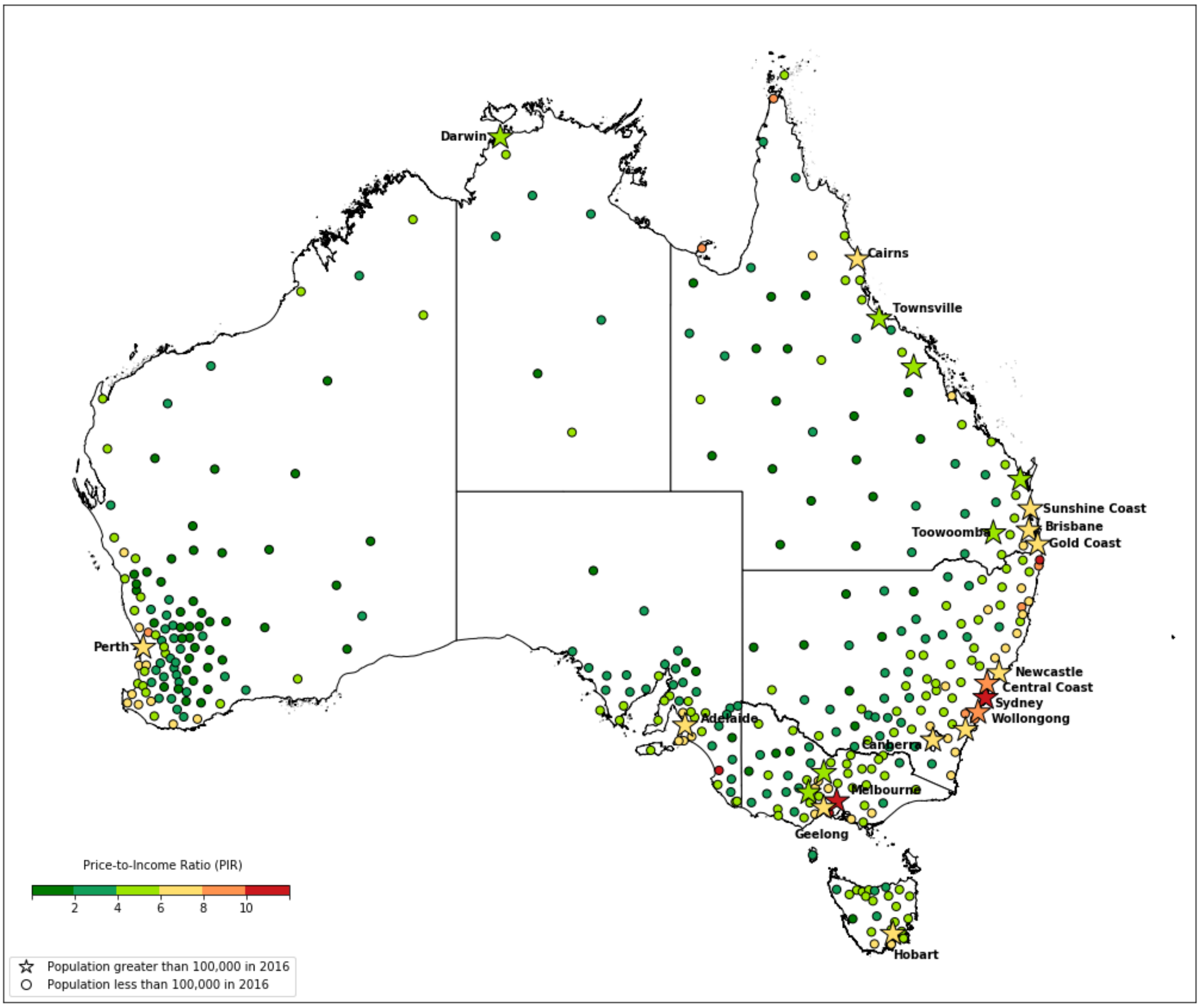

Figure 1. Price-to-Income Ratios Across Australian Urban Areas in 2016. Source: Income data corresponds to annual wages for those that are employed full-time, and has been sourced from the 2016 Census; and house sales value data denotes the average sales price of all recorded house transactions during 2016, and has been sourced from CoreLogic.

All variables, including employment and industry mix, are based on place of residence and not place of work. In general, larger cities tend to be younger and more educated. However, there are no clear spatial patterns with regards to the other control variables, such as labour force participation rates or industrial diversity.

With regards to industry mix, we have used a system of industry aggregation based on the 19-industry Australian and New Zealand Standard Industrial Classification (ANZSIC) system, the most widely accepted standard in the region for reporting industry-specific information. In some cases, we have aggregated industries based on similarities. For example, 'information media and telecommunications'; 'financial and insurance services'; and 'professional, scientific and technical services' have been aggregated under the label 'new economy services', based on their shared relationship with the 'knowledge economy'. In others, we have preserved individual ANZSIC categories, based on the size of the industry and its importance to regional and national products. For example, the mining sector has exerted significant influence on the Australian economy in recent years, and its effects need to be controlled for explicitly. Analogously, the large relative size of the health care and social assistance sector leads us to treat this sector independently as well. We treat the size of the accommodation and food services sector as a proxy for tourism-related economic activity in the urban area.

3. ECONOMETRIC FRAMEWORK

Our analysis involved time series data for the two measures of interest, namely wages and housing costs, over the years: 2001, 2006, 2011 and 2016. All data are available at the LGA level. However, multiple LGAs belonging to the same urban area were consolidated based on guidance from the Australian Statistical Geography Standard. We have 384 consolidated urban areas in our sample, and these comprise our unit of analysis throughout.

We analyse the time series data using the following model specification:

$$y_{lt} = \mu_t + \beta_1 p_{lt} + \beta_2 \ln p_{lt} + \gamma x_{lt} + \rho \sum_j w_{lj} f(y_{jt}) + \varepsilon_{lt} \qquad (1)$$

which represents a general structure similar to specifications used by other empirical studies of regional socioeconomic performance (Davies and Tonts, 2009; Bosker, 2007; Mankiw *et al.,* 1992). Over subsequent paragraphs, we describe each component of our specification in greater detail, and how it relates to the relevant literature.

The dependent variable of interest is denoted y_{lt}, and refers to wages or housing costs for urban areas l during year t. Separate (seemingly unrelated) regression equations were specified for each dependent variable. Some studies have used changes in these measures as dependent variables, as opposed to their absolute values, to control for differences in baseline values across regions (Bosker, 2007; Glaeser *et al.,* 1992). We tried both specifications, and we found models with absolute values as the dependent variables to perform uniformly better across different specifications.

The parameter μ_t denotes fixed effects specific to year t, and captures short-term deviations specific to particular years. For example, wages declined in 2009 across all regions nationwide due to the global financial crisis of the preceding year. The μ_t parameter would capture such effects specific to a particular year but applicable to all urban areas in the sample for that year.

The variable p_{lt} denotes the population of urban area l during year t. The effect of population size on the dependent variable may be linear, logarithmic, quadratic or some other functional form. The appropriate functional form was determined based on a comparison across goodness-of-fit and behavioural interpretation. For both dependent variables, the log-linear specification was the preferred specification. Correspondingly, the parameters β_1 and β_2 denote the correlation between the dependent variable and population size.

The variable x_{lt} denotes explanatory variables other than time and population that are hypothesized to have a significant effect on the dependent variables of interest. We listed these variables in Table 1, along with the sources for the data. Most of these variables capture major demographic differences across urban areas, as denoted by corresponding differences in age, education and employment.

Some studies have suggested including a standalone parameter α_l in equation (1), as a fixed effect specific to urban area l. The parameter α_l captures the influence of all other variables not included in our model that may vary across regions, and may have an effect on the dependent variable. The use of fixed effects to control for unobserved heterogeneity across different urban regions is the recommended approach in the literature, "as there are good reasons to believe (Islam, 1995) that the region-specific effects represent systematic differences in steady states that if left out of the estimation can bias the estimates of the parameters of interest" (Bosker, 2007). However, we only have four observations per urban area, which limited our ability to identify area-specific fixed effects, and consequently, we have excluded the parameter from our final specification.

Given the spatial nature of our data, and the positive spillover effects frequently received by the local economies of smaller urban areas situated in close proximity to major urban areas (Bosworth and Venhorst, 2018), we employ a spatial lag in our model specification, such that the value of the dependent variable in a particular urban area during a given year depends on the corresponding values of the same variable in surrounding urban areas during that same year. The parameter ρ captures the direction and magnitude of effect exerted by neighbouring areas, and the variable w_{lj} denotes the degree of connectivity between the urban areas l and j. There are many different ways in which w_{lj} might be constructed (for a comprehensive review of these different methods, see, for example, Anselin, 2013). In our case, we constructed distance-based weights based on the well-known gravity model, where the degree of connectivity between two urban areas is inversely proportional to the square of the physical distance separating these areas. We assumed further that urban areas that are more than 1000 km apart do not exert any influence on each other (i.e. $w_{lj} = 0$ for these areas l and j). The weights are normalized such that $\sum_j w_{lj} = 1$ for any urban area l.

Finally, the variable ε_{lt} is the residual term that captures the average effect of all other variables not included in our regression, and can be thought of as white noise. We assume that ε_{lt} is independently and identically distributed with variance σ^2 across all urban areas and time periods, where σ is a model parameter to be estimated. We are implicitly assuming that the noise in our data is not temporally correlated. In other words, we assume that random shocks at each time point in our data do not propagate to future time points. Relatedly, we are also assuming that the noise in our data is not spatially correlated, and that any spatial correlation is captured through the spatial lag variable. Our model specification allows ε_{lt} to be correlated across our two dependent variables of interest, for the same urban area and time period, resulting in a system of seemingly unrelated regression (SUR) equations.

All models were estimated through the PySal library (Rey and Anselin, 2010) in Python using an implementation of three-stage least squares estimation methods for a system of SUR equations with spatial lag. The estimation results for our final specifications are reported in Table 4. In the following section, we discuss these results in greater detail for each of the dependent variables.

4. ESTIMATION RESULTS

The estimation results for our seemingly unrelated and spatially lagged regression equations of wages and housing costs across 384 Australian urban areas over the years 2001, 2006, 2011 and 2016 are reported in Table 4. The model is able to explain roughly 78 per cent of the variation in annual full-time wages and roughly 71 per cent of the variation in housing costs across urban areas and time periods, indicating a high degree of explanatory power. Most of the parameters are consistent with prior expectations in terms of both the sign and the magnitude of their effect.

Both wages and housing costs grow log-linearly with the population of an urban area, indicating that larger urban areas have both higher wages and housing costs. The dynamics of income growth are such that the marginal increase in incomes is greater for smaller regional centres than it is for larger metropolitan areas. This is because the logarithmic coefficient dominates at smaller population sizes under 1 million, and the linear coefficient dominates at larger population sizes beyond 1 million. In the case of housing costs, the linear coefficient dominates the relationship, even at small population sizes, indicating that the marginal change in housing costs with population is roughly similar regardless of population size.

Ceteris paribus, average wages grow by roughly $1,500, and average housing costs grow by roughly $48,000, as the population of an urban area increases by 1 million. Stated differently, housing costs grow roughly thirty times faster with population size than wages, indicating that real wages do not keep up with the high costs of living associated with large cities. As an illustration, Figure 2 plots the PIRs for different population sizes, as predicted by our model for the case of Greater Newcastle and Greater Sydney. For the former, our model predicts that the ratio could rise from 7.0 to 8.4, if Greater Newcastle were to grow to the present size of Sydney. Similarly, for the latter, our model predicts that PIRs could rise from 13.6 to 14.8 by 2056, if Greater Sydney grows to its expected size of 9.2 million.

Next, we examine the effect of other variables on wages and housing costs. Remoteness has a strong positive effect on wages and a strong negative effect on housing costs. The former result reflects higher costs of living in more remote areas, where average wages must be commensurately higher to keep up with more expensive goods and services. And the latter result reflects deflated housing markets in more remote areas due to persistent out-migration to the larger cities. On average, our model finds that an urban area with a remoteness designation

Table 4. Estimation Results from the Regression Analysis. Source: the Authors.

Explanatory Variable	**Annual full-time wages ($)**		**Housing costs ($)**	
	est.	**p-value**	**est.**	**p-value**
Population				
Population (millions)	1,507	0.09	47,873	0.00
Natural logarithm of population	1,133	0.00	3,231	0.21
Population density (per sq. km)	1.08	0.76	-32.6	0.31
Geographic factors				
ASGS remoteness designation	3,088	0.00	-10,443	0.00
Spatial lag variable	0.065	0.10	0.379	0.00
Demographic factors				
Average age (years)	-1,114	0.00	-3,878	0.00
Proportion of population with a Bachelor's degree or higher (%)	1,293	0.00	12,395	0.00
Labour force participation rate (%)	28.0	0.07	-463	0.00
HHI index of industry mix	-8,243	0.14	291,904	0.00
Employment mix across industries				
Agriculture, forestry and fishing (%)	160	0.00	-1,704	0.00
Mining (%)	577	0.00	-1,323	0.00
Manufacturing (%)	-119	0.06	-1,370	0.01
Other 'blue collar' jobs (%)	500	0.00	4,645	0.00
New economy service (%)	-429	0.03	-213	0.90
Health care and social assistance (%)	-864	0.00	-2,547	0.33
Accommodation and food services (%)	3.64	0.97	2,444	0.30
Other services (%)	171	-	67.6	-
Time-specific fixed effects				
Year 2001	53,343	0.00	62,749	0.17
Year 2006	63,951	0.00	110,800	0.02
Year 2011	73,850	0.00	143,144	0.00
Year 2016	82,540	0.00	146,254	0.00
Correlation coefficients				
Annual full-time wages ($)	1.000	-	0.258	-
Housing costs ($)	0.258	-	1.000	-
Goodness of fit measures				
R-squared	0.778	-	0.708	-
Adjusted R-squared	0.775	-	0.705	-

of 5 (i.e. very remote) has average wages that are $12,000 higher and housing costs that are $40,000 lower than an urban area with a remoteness score of 1 (i.e. a major Australian city).

Spatial lag is small and statistically weak in the case of wages but large, positive and statistically significant in the case of housing costs. There are no wage benefits to urban areas situated in close proximity to a large metropolitan centre. However, these areas incur higher costs in terms of more expensive local housing markets, indicating that a rise in housing costs in an urban area can create inflationary pressures on housing markets in surrounding centres.

Surprisingly, population density is not found to have a statistically significant effect on either average wages or housing costs. Note that agglomeration economies should cause wages to rise with population density, as firms and workers locate closer to each other. Conversely, agglomeration diseconomies should cause housing costs to rise with population density, as more people are forced to share the same piece of land. However, our analysis finds that once we control for population size, remoteness and spatial lag, the effect of population density on both wages and housing costs is statistically insignificant.

Labour force participation rates have a positive though statistically weak effect on wages, and a negative and statistically significant effect on housing costs, but in both cases, the magnitude of the effect is very small: an increase by 1 per cent in the proportion of an urban area's population that works is found to increase the area's annual wages by $28 and decrease the area's housing costs by $463.

The effect of industrial diversity on wages is statistically insignificant, but the effect on housing costs is negative and significant. Urban areas that have higher HHI measures, i.e. lower employment diversity, have higher housing costs. Note that our employment data is based on place of residence and not place of work. Therefore, the HHI for an urban area measures employment diversity of local residents, which may be different from the diversity of local jobs in the urban area. Our finding likely reflects economic gentrification and segregation effects, where expensive local housing markets force residents working in low-income industries to leave, and the urban area itself becomes more homogenous over time (Kohn, 2013; Lees, 2008).

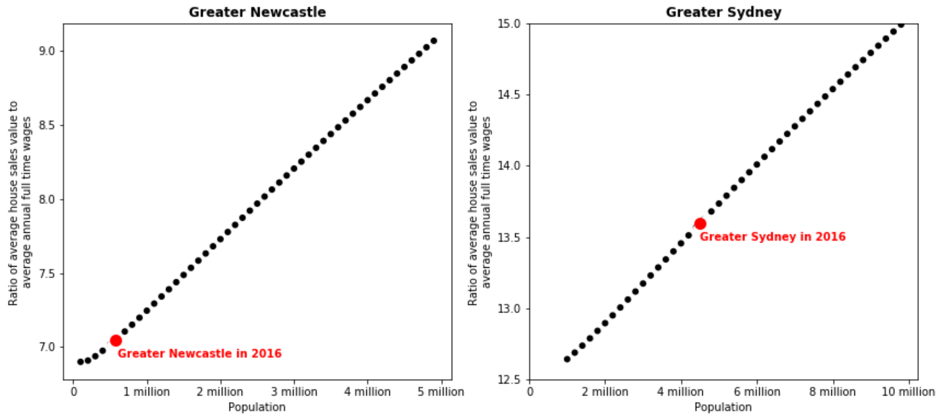

Figure 2. Predicted Ratio Between Average Home Sales Value and Average Full-Time Wages for the Greater Newcastle and Greater Sydney Urban Areas at Different Population Sizes. Source: the Authors.

In terms of employment mix, wages are found to be positively correlated with the proportion of residents employed in agriculture and mining, while housing costs are negatively correlated. Mining, and to a lesser extent agriculture, have been major engines for national economic growth (Downes *et al.,* 2014), and are expectedly associated with higher wages. However, both industries exert strong negative impacts on local environments, resulting in lower local housing costs. The proportion of residents employed in other 'blue collar' jobs is positively correlated with both wages and housing costs. Note that this measure includes residents employed in construction. The result is not surprising, as high local wages and housing costs can serve as incentives for increased construction activity. Conversely, the proportion of residents employed in health care and social assistance has the strongest negative effect on both wages and housing costs: an increase in the proportion of an urban area's labour force that works in the sector by 1 per cent is found to decrease the area's average annual wages by $864 and the average housing costs by $2,547. In general, a large 'other blue collar' sector (which includes construction-related economic activity) is an indicator of a strong local economy, and a large 'health care and social assistance' sector is an indicator of a weak local economy.

Finally, average wages are positively correlated with housing costs. In other words, urban areas with high average wages expectedly have higher housing costs.

5. CONCLUSIONS

Much has been written about the recent scarcity of affordable housing in Australia. Dwelling PIRs have nearly doubled between 1980 and 2018 (OECD, 2019), and mortgage debt has increased from 32% of household disposable income and 20% of GDP in 1990, to 142% and 88% in 2010, respectively (MacKillop, 2013). The situation is particularly grim in the largest cities. For example, average housing costs in Sydney and Melbourne are between two and three times higher than those in other Australian cities and regional centres (Chappell and Campbell, 2018).

Between 2016 and 2056, the Australian population is projected to increase by 18.3 million people, and approximately 56 per cent of this growth is expected to occur in Australia's two biggest cities – Sydney and Melbourne. It is commonly believed that bigger cities are more economically efficient than smaller cities because of improved access to labour and supporting services, knowledge spillovers between firms, and the creation of a local market. However, increased agglomeration might also lead to greater urban costs in the form of housing unaffordability, and highly concentrated patterns of population growth could further exacerbate the growing divergence between wages and housing costs in some of Australia's largest cities.

This study sought to examine the relationship between an urban area's population size, wages and housing costs, while controlling for other confounding factors, such as differences in the local demography and the local industry mix. Using nationwide LGA-level data over the time period 2001-16, collated from the ABS and other sources, we ran a series of multivariate regressions to understand these relationships better. Our analysis revealed that wages and housing costs rise monotonically with the population of an urban area. However, on average, housing costs rise roughly thirty times faster with population size than wages, indicating that real wages do not keep up with the high costs of living in big cities. For example, our model predicts that PIRs in Greater Newcastle could rise from 7.0 to 8.4, if the city grows to the size of Sydney. And PIRs in Sydney themselves could rise from 13.6 to 14.8 by 2056, if the city grows to its expected size of 9.2 million.

As mentioned previously, the limited availability of affordable housing in Australia has been attributed to a variety of supply and demand-side factors that include housing tax policy, land-use planning and zoning laws, foreign investment in real estate, declining government support for social housing, changes in residential preferences, etc. (Lawson *et al.*, 2019;

Eccleston *et al.*, 2018; Gurran *et al.*, 2018). Our analysis confirms that spatially unbalanced patterns of regional growth that have concentrated more people in fewer cities have also contributed to the situation. National and state governments in Australia have frequently adopted the view that significant positive agglomeration economies will grow the biggest cities even bigger, and any urban costs can potentially be offset through improved planning and management of housing markets, infrastructure systems, and other supporting services. Findings from our analysis serve as a cautionary warning against the pursuit of such regional growth strategies.

REFERENCES

Australian Bureau of Statistics (2008). Population Projections, Australia, 2006 to 2101. Cat. no. 3222.0. Canberra: Australian Bureau of Statistics.

Adamy, J. and Overberg, P. (2019). Millennials continue to leave big cities. *The Wall Street Journal*, Sep 26, 2019.

Anenberg, E. (2016). Information frictions and housing market dynamics. *International Economic Review*, *57*(4), pp. 1449-1479.

Anselin, L. (2013). *Spatial econometrics: Methods and models (Vol. 4)*. Berlin: Springer Science & Business Media.

Bosker, M. (2007). Growth, agglomeration and convergence: a space-time analysis for European regions. *Spatial Economic Analysis, 2*(1), pp. 91-100.

Bosworth, G. and Venhorst, V. (2018). Economic linkages between urban and rural regions–what's in it for the rural? *Regional Studies*, *52*(8), pp. 1075-1085.

Chappell, J. and Campbell, N. (2018). The Housing Gap - Sydney, Australia. In W. L. Filho (Ed). *Sustainable Development Research in the Asia-Pacific Region: Education, Cities, Infrastructure and Buildings*, Berlin: Springer.

Daniel, L., Baker, E. and Lester, L. (2018). Measuring housing affordability stress: Can deprivation capture risk made real? *Urban Policy and Research*, *36*(3), pp. 271-286.

Davies, A. and Tonts, M. (2010). Economic diversity and regional socioeconomic performance: An empirical analysis of the Western Australian grain belt. *Geographical Research, 48*(3), pp. 223-234.

Downes, P. M., Hanslow, K. and Tulip, P. (2014). The effect of the mining boom on the Australian economy. *Reserve Bank of Australia research discussion paper*, (2014-08).

Duranton, G. and Puga, D. (2004). Micro-foundations of urban agglomeration economies. In *Handbook of regional and urban economics* (Vol. 4, pp. 2063-2117). Amsterdam: Elsevier.

Eccleston, R., Verdouw, J., Flanagan, K., Warren, N., Duncan, A., Ong, R., Whelan, S. and Atalay, K. (2018). Pathways to housing tax reform. AHURI Final Report No. 301. Melbourne: Australian Housing and Urban Research Institute Limited.

Florida, R., Adler, P. and Mellander, C. (2017). The city as innovation machine. *Regional Studies*, *51*(1), pp. 86-96.

Fujita, M., Krugman, P. R. and Venables, A. J. (2001). *The spatial economy: Cities, regions, and international trade*. Cambridge: MIT Press.

Fujita, M. (1989). *Urban economic theory: land use and city size*. Cambridge: Cambridge University Press.

Glaeser, E. L., Kallal, H. D., Scheinkman, J. A. and Shleifer, A. (1992). Growth in cities. *Journal of political economy*, *100*(6), pp. 1126-1152.

Glaeser, E. L. (1998). Are cities dying? *Journal of economic perspectives*, *12*(2), pp. 139-160.

Glaeser, E. (2011). *Triumph of the City*. Sydney: Pan Macmillan Australia.

Glaeser, E. L. and Mare, D. C. (2001). Cities and skills. *Journal of labor economics*, *19*(2), pp. 316-342.

Glaeser, E. L., Gottlieb, J. D. and Ziv, O. (2016). Unhappy cities. *Journal of labor economics*, *34*(S2), pp. S129-S182.

Gurran, N., Rowley, S., Milligan, V., Randolph, B., Phibbs, P., Gilbert, C., James, A., Troy, L. and van den Nouwelant, R. (2018). Inquiry into increasing affordable housing supply: Evidence-based principles and strategies for Australian policy and practice. AHURI Final Report 300. Melbourne: Australian Housing and Urban Research Institute Limited.

Henderson, J. V. (1974). The sizes and types of cities. *The American Economic Review*, pp. 640-656.

Islam, N. (1995). Growth empirics: A panel data approach. *The Quarterly Journal of Economics*, *110*(4), pp. 1127-1170.

Jacobs, J. (1969). *The economy of cities*. New York: Random House.

Kaufman, B. E. (2007). The impossibility of a perfectly competitive labour market. *Cambridge Journal of Economics*, *31*(5), pp. 775-787.

Kohn, M. (2013). What is wrong with gentrification? *Urban Research & Practice*, *6*(3), pp. 297-310.

Krugman, P. (1991). Increasing returns and economic geography. *Journal of Political Economy*, *99*(3), pp. 483-499.

Lawson, J., Denham, T., Dodson, D., Flanagan, K., Jacobs, K., Martin, C., Van den Nouwelant, R., Pawson, H. and Troy, L. (2019). Social housing as infrastructure: rationale, prioritisation and investment pathway. AHURI Final Report No. 315. Melbourne: Australian Housing and Urban Research Institute Limited.

Lees, L. (2008). Gentrification and social mixing: towards an inclusive urban renaissance? *Urban Studies*, *45*(12), pp. 2449-2470.

MacKillop, F. (2013). Sustainable as a basis of affordable? Understanding the affordability ‘crisis’ in Australian housing. *Australian Planner*, *50*(1), pp. 2-12.

Mankiw, N. G., Romer, D. and Weil, D. N. (1992). A contribution to the empirics of economic growth. *The Quarterly Journal of Economics*, *107*(2), pp. 407-437.

Marshall, A. (1890). *Principles of economics*. London: Macmillan.

McGuirk, P. and Argent, N. (2011). Population growth and change: Implications for Australia's cities and regions. *Geographical Research*, *49*(3), pp. 317-335.

Mills, E. S. (1967). An aggregative model of resource allocation in a metropolitan area. *The American Economic Review*, *57*(2), pp. 197-210.

Organisation for Economic Co-operation and Development (2019). Housing prices. Online version accessed September 2019, https://data.oecd.org/price/housing-prices.htm.

Paris, C. (2010). *Affluence, mobility and second home ownership*. London: Routledge.

Quigley, J. M. (1998). Urban diversity and economic growth. *Journal of Economic Perspectives*, *12*(2), pp. 127-138.

Rey, S. J. and Anselin, L. (2010). PySAL: A Python library of spatial analytical methods. In M. M. Fischer and A. Getis (Eds). *Handbook of applied spatial analysis* (pp. 175-193). Berlin: Springer.

Richardson, H. W. (1995). Economies and diseconomies of agglomeration. In H. Giersch (Ed). *Urban agglomeration and economic growth* (pp. 123-155). Berlin: Springer.

Rosenthal, S. S. and Strange, W. C. (2004). Evidence on the nature and sources of agglomeration economies. In V. Henderson and J. F. Thisse (Eds). *Handbook of regional and urban economics* (Vol. 4, pp. 2119-2171). Amsterdam: Elsevier.

Valadkhani, A., Smyth, R. and Worthington, A. (2017). Regional seasonality in Australian house and apartment price returns. *Regional Studies*, *51*(10), pp. 1553-1567.