ARC Centre of Excellence in Population Ageing Research

Working Paper 2018/23

The Income Elasticity of Housing Demand in New South Wales, Australia

Xiangling Liu

The Income Elasticity of Housing Demand in New South Wales, Australia

Xiangling Liu
xiang.liu@unsw.edu.au
CEPAR & School of Economics, UNSW
Accepted version by "Regional Science and Urban Econoimcs"

January 9, 2019

Abstract

This paper studies the theoretical relationship between house prices and income by using the user cost equilibrium condition. Empirically, the long-run and short-run dynamics of this relationship are studied by using data for 144 Local government areas (LGA) over 25 years from 1991 to 2015 in the state of New South Wales, Australia. The income elasticity of house prices for the state is estimated to be 1.07 by multi-factor panel data models and the cointegration analysis. The income elasticities across locations demonstrate a spatial pattern, higher in Sydney and the Sydney surrounds and diminishing as going to inland regional and rural areas. The Granger Causality of the co-integrated relationship has been studied sequentially and proves the unidirectional causality from income to house prices. Finally, the state-wide common factors are found to show widespread significance in contrast to the Sydney-wide common factors which only impact significantly the areas that surround Sydney within a certain spatial range.
JEL classification: R12, R31, C33
Keywords: Income elasticity; cross-section dependence; house price dynamics; cointegration

1 Introduction

Australia's housing market has experienced a steady level of growth for more than a quarter century, with only a few transitory periods of slowing growth during episodes of modest or stagnant economic growth. Housing investment and consumption depend on the disposable income, which impacts the household wealth and further influences borrowing and spending decisions. The link between house prices and income has been traditionally quantified by studying the income elasticity of housing demand. The accurate estimate of it is of much concern for gauging the economic health and deliberating on macroeconomic policy. However, the long-term growth in the housing market poses some challenge to obtain unbiased econometric estimates for economic parameters, since the contrasts created by the ups and downs of different economic time series are missing. Housing markets may be better characterized as a series of interconnected subregional markets which enable to incorporate cross-sectional information to help identify the parameters. To facilitate such a purpose, the analysis in this paper uses data for the 144 local housing markets in the state of New South Wales in Australia over 25 years from 1991 to 2015 for studying the house price dynamics in terms of income and other fundamental economic variables.

One key element that needs to be taken into account for studying the local housing markets is that developments of local housing markets are likely to be impacted by spatial or non-spatial diffusion effects. Spatial patterns of the distribution in house prices are considered to arise when cross section units are subject to common effects (Meen 1999; Holly, Pesaran, and Yamagata 2010) and (or) if there are endogenous spatial effects caused by spatial or other forms of local dependencies (Alexander and Barrow 1994; Ashworth and Parker 1997 and Cook 2003). Meen suggests that common factors, mainly described as macroeconomic conditions, including the observed changes in interest rates and oil prices, or the unobserved changes, such as technological changes, can result in the heterogeneity in the responses of sub-markets in a given geographical area to the overall state of the macro-economy. Alexander and Borrow also suggest that interaction effects, for example, inter-regional interactions caused by different preferences to coastal amenities, distances to CBD and culture, and the location contiguousness can also naturally lead to heterogeneities in house prices across locations.

The potential bias in the income elasticity of housing demand derived from time series or cross-sectional estimates is well recognized to yield higher estimates. One would expect estimation bias due to the insufficient considerations to spatial effects that are caused by the spatial house price variations. For Australia, Abelson (1994), Bourassa and Hendershott (1995), and Abelson and Chung (2005) use time series approach to study the boom and the stagnation pattern of house prices for the capital cities in Australia and obtain the income elasticities between 1.3-1.7. One possible problem is that time dummies and fixed effects they use to control the regional effects overlooked the potential endogenous problems caused by the interactions across the cities. In the U.S housing markets, Gallin (2006) estimates the income elasticity of housing demand in a panel of 95 metro areas over 23 years between 1.45 and 1.71 yet without confirmation for the presence of the co-integrated relationship between them. In contrast, by explicitly taking account into the cross-section dependence and heterogeneity, Holly, Pesaran, and Yamagata (2010) finds the income elasticity to be a unit in their panel data across U.S 46 states over 29 years. In Spain, Fernández-Kranz and Hon (2006) estimate the income elasticity of housing demand in fifty Spanish provinces from 1996 to 2002 between 0.70 and 0.95 in comparison to the estimate of 2.8 by long-run equilibrium models using time-series methods (Maza and Pages 2007).

In this paper, panel data models are estimated for house prices using the common correlated effects (CCE) estimators (Pesaran 2006) which allow the unobserved common factors to be (possibly) correlated with the exogenously given individual-specific regressors and the factor loadings to differ over the cross-section units. One advantage of this approach is that when the cross-section dimension has a relatively larger size than the time dimension, means of (weighted) cross-section aggregates could be augmented in the panel regressions to filter out the differential effects of unobserved common factors. This approach also provides consistent estimates for slope coefficients despite heterogeneity and cross-sectional dependence, or a presence of a spatial error process. In comparison, approaches in spatial econometrics deal with spatial effects by constructing the $N \times N$ spatial weight matrix where each cross-section unit is assumed to relate to its neighbor(s) by some form, such as Anselin (1988) and Baltagi, Song, and Koh (2003), LeSage and Pace (2010), Kapoor, Kelejian, and Prucha (2007), and Anselin, Le Gallo, and Jayet (2008). These approaches, however, require appropriate designs of the spatial matrices, the misspecification of which may lead to substantial size distortions in the tests based on ML or Quasi-ML estimators (Pesaran and Tosetti 2011).

In the empirical analysis, the income elasticity of housing demand is estimated to be 1.07 which verifies a long-run cointegrating relationship between real house prices and real income despite short-run dynamics. Fluctuations in this relationship appear to be the most volatile in the regional NSW and relatively stable in the Sydney and Sydney Surrounds. The income elasticities across the locations demonstrate a spatial pattern, which is the highest in the Sydney (metropolitan), less significant in its surroundings and the coastal-line areas, and further diminishing as going to the inland regional areas.

With the confirmation of the cointegrating relationship, the Granger causality of the relationship has also been studied and the causality is found to be unidirectional from real income to real house prices, where interestingly, the short-run shocks are also found only significant from the real income to the real house prices. The results suggest the within-sample Granger exogeneity of real house prices to the real income, and the within-sample Granger endogeneity of real income to the real house prices. It seems that when real house prices are subject to some either transitory or permanent shocks, the changes in the real income act to be the brunt in the adjustment of clearing the disequilibrium.

Another finding is that the factor loadings on the state-wide common factor are statistically significant (more than 0.7) in the most of coastal areas of the state, including Sydney, the most of the Sydney Surrounds, the Hunter region, the Mid North Coast region, the Richmond-Tweed region and the most of the South-Eastern region. In contrast, the significance of the loadings on the Sydney-wide common shocks demonstrates another interesting spatial pattern - more than 0.7 in the Sydney and the LGAs within the commuting distances of Sydney, that is, the north of Sydney as far as to the Port Stephens, the South of Sydney as far as to the Eurobodalla, and the west of Sydney as far as to the Centre-West region. It appears that the state-wide common shocks have been the common thread that binds all of the LGAs in the state, and the Sydney-wide common shocks have propagated but only to a certain spatial range surrounding the Sydney.

The remaining of this paper has the following structure. It firstly provides a theoretical relationship between real house prices and real income. The econometric methods used in this paper are then described, followed by a description of the panel dataset and a preliminary data analysis. The empirical results, including unit root tests, income elasticities, cointegration analysis and factor loadings are reported sequentially. The last section concludes.

2 Model House Prices

Literature has conventionally utilized the user cost equilibrium condition for the pricing of housing prices which suggests that the annual cost of owning of a residential property should equal to the one year rent (Poterba 1984; Himmelberg, Mayer, and Sinai 2005; Hill and Syed 2016). The equilibrium condition also indicates in an efficient market, a prospective renter must be indifferent between renting and owning. The user cost of owning can be formulated as,

$$U_t = (r\delta_t + r_t^d)P_t - E_t(P_{t+1} - P_t|\xi_t) + \Re P_t + F_t \tag{1}$$

where P_t is the market price of the property, δ_t is the loan to value ratio, r the real mortgage interest rate, r_t^d is the annualized depreciation rate of the building structure in relation to the house price. $E_t(P_{t+1} - P_t|\xi_t)$ is the offset annualized capital gain where ξ_t is the information set and $\Re$ is the rate of risk premium (as opposed to investing in financial assets). F_t accounts for the trading cost and the opportunity cost of the forgone interest that an investor could have earned by investing in the financial asset market. F_t

is written as,

$$F_t = (1 - \delta_t + \ell)P_t * r^f \tag{2}$$

where ℓ is the annualized rate of the trading cost of the investment in housing and r^f is the risk-free rate of return for lending. By assuming the lending rate equal to the borrowing rate, $r = r^f$, the capital income R_t of a property in the user-cost equilibrium should be,

$$R_t = ((1 + \ell)r + r_t^d + \Re)P_t - E_t(P_{t+1} - P_t|\xi_t) \tag{3}$$

equivalently, the equation for house prices is written as,

$$P_t = \frac{1}{1 + r_t^d + \Re + (1 + \ell)r}[R_t + E_t(P_{t+1}|\xi_t)] \tag{4}$$

Assuming housing must be purchased through the acquisition of a mortgage loan that households can borrow against a portion of the house value and their future income. The maximum mortgage allowance that a household can borrow given the household taxable income Y_t is a function of the household real disposal income, the real interest rate of borrowing, and the term of mortgage[1]. The maximum mortgage allowance of the household can be written as,

$$B_t = \kappa Y_t \left(\frac{1 - (1 + r)^{-\tau}}{r}\right) \tag{5}$$

where κ is the proportion of household disposable real income going to mortgage repayments, and τ is the term of mortgage. Assuming the household uses a portion c of the mortgage allowance to purchase a property with the maximum loan to value ratio $\bar{\delta}$, the purchased property price is written,

$$P_t = \frac{cB_t}{\bar{\delta}} \tag{6}$$

At time t, the net cost of the real housing services is,

$$U_t' = (r(1 + \ell) + r_t^d)P_t \tag{7}$$

[1]The relationship between income, interest rates, and the typical amount of a mortgage offered by a financial institution is generally based on the present value of an annuity. The annuity is the fraction of current disposable income κY_t that goes toward mortgage repayments and is discounted at the current mortgage interest rate for a horizon equal to the term of the mortgage τ.

Substitute the equations (4-6) to (7),

$$U_t' = \alpha_t Y_t \tag{8}$$

where $\alpha_t = \frac{c}{\delta}\kappa(1 + \ell + \frac{r_t^d}{r})\big(1 - (1+r)^{-\tau}\big)$, $0 < \alpha_t < 1$ converges to a constant and is assumed as a stationary process. The user-cost equilibrium condition for an efficient market implies renters are indifferent from owning and renting for one period. The one year rent of a renter equal to the one year net cost of an owner,

$$R_t = U_t' \tag{9}$$

Substitute the equation (9) to (4), the house prices can then be obtained as the sum of future net costs of housing services discounted at the current expected gross rate of return in housing markets.

$$P_t = \sum_{j=0}^{\infty}\left(\frac{1}{1 + r_t^d + \Re + (1+\ell)r}\right)^j E(\alpha_{t+j}Y_{t+j}|\xi_t) \tag{10}$$

which can be written equivalently as,

$$\frac{P_t}{Y_t} = \sum_{j=0}^{\infty} E\left(\alpha_{t+j}\prod_{s=1}^{j}\left(\frac{1 + g_{t+s}}{1 + r_t^d + \Re + (1+\ell)r}|\xi_t\right)\right) \tag{11}$$

where $g_t = \Delta lnY_t$ is the growth rate of real disposable income. By making general assumptions, the expression above shows that if the processes generating α_{t+j} and g_{t+j} are stationary, the price-income ratio will be stationary and lnP_t and lnY_t will be cointegrated with the cointegrating vector (1,-1).

3 Econometric Model

The previous section models the relationship between the real house prices and the real disposable income. The following econometric factor model is employed to empirically study this relationship,

$$p_{it} = a_i + \beta_y y_{it} + \beta_q \Delta q_{i,t-1} + \beta_r r_t + \boldsymbol{\hbar}'\boldsymbol{f}_t + e_{it}, \tag{12}$$

where $i = 1, 2, .., 144$; i represents the ith local housing market and it is the housing market for the ith local government area (LGA). t=$1, 2, .., 25$ indicates the time; p_{it} is the logarithm of real house prices in the ith LGA during time t, y_{it} is the logarithm of real income per taxpayer in the ith LGA during time t, and β_y is the income elasticity of house prices measuring the percentage change in the demand for houses in response to a one percent increase in income. Based on the theory developed in the last section, p_{it} is likely to be cointegrated with y_{it}, and the β_y is expected to be a unit, that is, the p_{it} responds to y_{it} with the unit elasticity. r_t is the rate of borrowing that represents the cost of housing services. $\Delta q_{i,t-1}$ controls the short run dynamics where Δ is the difference operator and $\Delta q_{i,t-1}$ is the first-lag of the population growth on the ith cross-section unit at time t. $\boldsymbol{f}_t$ is assumed to be an m-dimensional vector of unobservable exogenous common factors[2], which might influence house prices across regions of NSW in some manner, such as the employment rates, policy effects, expectations, or technologies. In practice, $\boldsymbol{f}_t$ are proxied by the cross-section averages of p_{it}, y_{it} and $\Delta q_{i,t-1}$, written as $\bar{p}_t, \bar{y}_t$ and $\bar{q}_t$ respectively. $\boldsymbol{\hbar}_i'$ is a $1 \times m$ vector of heterogeneous factor loadings. The idiosyncratic error term, e_{it}, is assumed to allow for the weak spatial dependence and be independent of the exogenous independent variables and $\boldsymbol{f}_t$.

It is worth to note that the estimation for the above model (12) focuses on the state-wide average of the individual estimates for LGAs. One issue that needs to be pointed out is that the above model treats all the cross section units as equally important in contributing to the average estimate. However, LGAs may not equally contribute to the importance of the income elasticity of the state. For example, it is highly likely that transactions are higher in the Sydney LGAs than in the LGAs in the rural areas of the state. Accounting for the likely weights of transactions in LGAs rather than treating them as all equally important would help to give a more reasonable estimate. To deal with the matter, I consider approximating the weight that one LGA contributes to the income elasticity of the state by the weight of the population that the LGA has. Accordingly, the logarithm of population $q_{i,t}$ is added to the model in equation (12) for alternative estimations.

The model (12) nests a number of panel data specifications. The simpler panel models do not consider the unobserved common factors ($\boldsymbol{\hbar}_i = 0$). If the slope coefficients are homogeneous, $\boldsymbol{\beta}_i' = \boldsymbol{\beta}'$, then (12) collapses to a standard fixed-effects model (possibly with spatially correlated errors), otherwise if allowing for heterogeneity in the slope coefficients, (12) yields the dynamic heterogeneous panel model without considering the cross-section dependence. The general version of the model (12) takes into account the unobserved common factors in the panel model, which allow for cross-section dependence in the error term. In the empirical analysis, these models are estimated sequentially for comparative analyses. CCEP and CCEMG estimators by Pesaran (2006) are used for obtaining the estimates. FE and MG estimators are used for comparisons.

In the following empirical analysis, cross-section dependence tests and panel unit root tests are conducted to test for the cross-section dependence and the stationarity in the

[2] In this study, m is assumed to be equal to the number of included exogenous regressors.

panel variables or in the error term of the econometric models throughout the econometric analyses. In terms of the cross-section dependence test, there are various tests for cross-section dependence in the panel data for the case of (small) T and (large) N (Frees 1995; Sarafidis, Yamagata, and Robertson 2009). In this study, the cross-section dependence (CD) test due to Pesaran (2004) is used. One advantage of the CD test is that it can be applied to a wide variety of models, including heterogeneous dynamic models with multiple breaks and non-stationary dynamic models with small/large N and T.

The earlier panel unit root tests are built on the assumption of cross-section independence such as (Harris and Tzavalis 1999; Maddala and Wu 1999; Hadri 2000; Choi 2001 and Levin, Lin, and James Chu 2002). Choi (2001) finds in a Monte Carlo analysis that a test based on the inverse normal distribution has the best trade-off in terms of size and power for panels with a large size for N. In comparison, the panel unit root tests developed recently (Bai and Ng 2004; Moon and Perron 2004; Pesaran 2007, and Pesaran, Vanessa Smith, and Yamagata 2013) have relaxed this assumption of the cross-section independence. The CIPS test by Pesaran (2007) and the CIPSM test by Pesaran, Vanessa Smith, and Yamagata (2013) control for cross-section dependence by using factor error specification models and also allow for serial correlations in the idiosyncratic components. One advantage of CIPS (CIPSM) is that unlike the principal component methods by Bai and Ng (2004) and Moon and Perron (2004), CIPS does not require estimating the number of unobserved common factors for obtaining valid individual CADF statistics. In addition, CIPS statistics are based on the simple averages of the individual CADF statistics and are asymptotically consistent as long as the number of individual-specific variables is greater than the true number of common factors. In this study, the traditional panel unit root test by Choi (2001), and the recently developed CIPS and CIPSM tests are applied to panel unit root tests.

4 Data

In literature, arguments arise on the measurement problems of housing consumption and income (such as De Leeuw 1971 and Wilkinson 1973), since the scope of individual investigations differs from various factors, such as socio-demographic variables, housing qualities and locations. To achieve a conformity, an important distinction is whether the transitory components of income and house prices are excluded. Slater (1986) suggests that grouping of data is a familiar means of 'washing out' transitory variations in the measured income. Lee (1968) uses a measure of 'permanent' income, in particular, makes an extensive use of socio-demographic variables such as the family size, age, and the socio-economic class of the head of the household, the inclusion of which he argues reduces the chance of overestimating the income elasticity of demand. Yet the longitudinal measures of these individual factors are usually difficult and even harder are the longitudinal measures of these factors across regions. As a result, some aggregation methods have to be used. By assuming all households have the same propensity to consume in terms of housing services, the aggregated method of 'average or median' is considered as an appropriate estimation for grouping the data at some geographical level for empirical analyses.

One concern with using the median house prices as the estimates for house prices in each LGA rather than the micro-level housing data is that the quality of the median purchased dwellings differs significantly across LGAs. For example, by using cross-sectional transactional housing data over 2001-11, Hill and Scholz (2017) hedonically measure house prices across space and time and find evidence of a gradual improvement over time in the quality of transacted dwellings within postcodes in Sydney. The improvements are primarily locational -by postcodes. However, Hill and Scholz also find that the extra precision provided by geospatial data as compared with postcode dummies has only a marginal impact on the resulting hedonic price index. As a result, they advise that the postcode dummies are sufficient to control for locational effects in a hedonic house price index. Analogously, the regional effects in the hedonic price indices at a higher geographical level will also be considered to be captured by the regional dummies. In the current panel study, one advantage is that the panel data analysis will allow the regional fixed effects to be captured by the cross-sectional dummies. In addition, if the variables are stochastic across the cross-section units and contain unit roots, the estimation will be considered as valid if the existence of one (or more) cointegrating relationships could be proved. In the econometric analysis in the following sections, the econometric model will appropriately control the fixed regional effects, the panel unit root tests will test for the integration of the variables of interest, and the cointegration analysis will provide evidence of the presence of the cointegrating relationship.

The panel dataset in this analysis consists of 144 LGAs in New South Wales with data observed annually for the period 1991 to 2015. Residential property prices for each LGA are measured using median prices obtained from the New South Wales Department of Housings Sales Reports. Quarterly observations are available for the most of LGAs from March 1991 to June 2016, classified by the "Strata" and the "Non-strata" properties and the classification of the "non-strata" properties is used for this analysis. The LGAs Balranald, Brewarrina, Central Darling, Conargo, Jerilderie, Murrumbidgee, Urana, Gwydir are dropped from analysis due to more than half of the observations for the price data are missing. To obtain the annual observations, the quarterly observations of house prices in a financial year is averaged, yielding 25 observations for houses in each LGA.

Estimates of income for LGAs are based on data from the Australian Taxation Offices Taxation Statistics (ATO). For the period 1990-91 to 2005-06, the Bureau of Infrastructure, Transport and Regional Economics (BITRE) reports the real income per taxpayer (in 2007-08 prices) by LGAs (using figures derived from Australian Taxation Office data and the consumer price index (CPI) for Australia). For the period 2006-07 to 2010-13, the Australian Bureau of Statistics (ABS) reports data by LGAs for Estimates of the average total Personal Income (excl. Government pensions & allowances), Time Series, 2005-06 to 2010-13. These figures are converted to constant 2007-08 prices using the CPI for Australia. There is no comparable income data by LGAs for the financial years 2013-14 and 2014-15, to cope with this problem, the growth rate of real income per taxpayer in these two years is assumed to be equal to the growth rate for 2011-13.

Data for the population are obtained from a number of sources. The NSW Local Grants Commission produces the data for population by LGAs over the period 1990-91 to 2000-01, with the data for the financial year 1991-92 missing. Without loss of generality, this missing data are imputed by using the average of the data in the financial years 1990-91 and 1992-93. For the remaining financial years 2001-15, ABS reports the Dataset: ERP by LGA (ASGS 2015), 2001 to 2015 for population by LGAs.

Two additional variables in this study are the CPI for Sydney and the real interest rate, which use the Australian Government indexed bond with the longest maturity from the RBA spreadsheet Capital Market Yields (F2). To conform with the annual frequency, both of the variables use the financial year average data. The collected data for house prices are divided by the CPI for Sydney to transform into the real values.

5 Preliminary Data Analysis

In the following empirical analysis, symbols for variables and their explanations are tabulated in Table 1. The urban theories imply that the concentrating economic activities in cities lead to agglomeration economies, and the existing central business districts (CBD) must be more intensified with the urban fringe extending away from the CBD in the urban equilibrium (Leunig and Overman 2008 and Liu and Otto 2017). In order to give some indication if there is the presence of such urban development patterns, the LGAs are grouped into three areas for comparison analysis, one for the LGAs within the Sydney (metropolitan), denoted as "Sydney", one for the LGAs surrounding the Sydney (metropolitan), denoted as "Sydney Surrounds", and one for the rest of the state, denoted as "Regional/Rural NSW". Figure 1 shows the position of the three LGA-groupings and the regions of the state. Table 2 lists the LGAs within each region and indicates the grouping that each LGA belongs. Figures 2-4 make the time series plots for the average of the changes of log real house prices and log real income respectively, and the average of the ratio of log real house prices to log real income by LGA-groupings.

Figures 2-3 demonstrate that fluctuations over time in either the real house prices or the real income are similar across the three classified areas. Real house prices have shown strong growth over the last a quarter century with a few periods of mild transitory declines (about 5%), 1995-1996 and 2005-2009[3]. The 2002-2004 witnessed dramatic house prices growth. Rather than Sydney grew the most, the Sydney Surrounds had increases about 20% and the Regional/Rural NSW 15%, in comparison with 10% in Sydney.

In comparison, there were more fluctuations in the real income during the period. The real income dropped in the early and the mid 1990s, 2000, 2002-2004 and 2008. During 2002-2004 when there were house price booms, the real income had a continuous fall rather than had a growth in the same period. One outstanding point is that the real income had an extraordinary growth in the Regional/Rural NSW in 2011 when 25 out of 101 Regional/Rural LGAs had the growth rate more than 15% and the other LGAs in the area had 11% growth on average, in comparison with the 4% in Sydney and the 5% in the Sydney surroundings. Since 2011, the real income has maintained a relatively stable growth rate at about 2% for the rural and regional area, 3% for Sydney and the 4% for the Sydney surrounds.

On average, Figure 4 shows that the ratio of real house prices to real income in Sydney was relatively stable, from 1.92 to 1.24, much higher than that in the Sydney surrounds and the rural and regional areas where the ratio seem had more change over the period. During the 2002-2004, the price-income ratio had common dramatic increases across all the areas, where the Sydney surrounds appeared to have the most significant change, from 1.16 to 1.22 catching up the level in Sydney, but after which it was dropping up to the recent. It is also worth to noting that the price to income ratio in Sydney seems to be growing and drifting further from the other areas in the state in the recent years.

6 Empirical Results

6.1 Preliminary econometric tests

6.1.1 Panel unit root tests

A necessary but not sufficient condition for cointegration is that for the variables in the model which contain unit roots, each of the variables should be integrated of the same

[3]Abelson (1994) noted that real house prices in Sydney fell by 17% in 1990 and real house prices moved in cycles over the period 1970-1990.

order. In order to verify the existence of this preliminary condition, before the formal empirical analysis for the econometric model, panel unit root tests via various testing procedures are carried out to investigate the unit roots in the data.

Table 4 reports the results of panel unit root tests for the panel variables in the dataset. These include the Z-statistic recommended by Choi (2001) that does not account for cross-section dependence in the panel units, and the $\overline{CADF}$ statistics for both CIPS and CIPSM tests which allow for the cross-section dependence.

In the upper half of the table, ADF regressions include a constant and allow for one or two lags of the dependent variables respectively. In the bottom half of the table, ADF regressions include both a constant and a time trend. For the variables, p_{it}, y_{it}, q_{it}, the computed Z-statistics suggest that the null hypothesis of a unit root cannot be rejected, except in one case for real income. In contrast for the differenced panel variables, Δp_{it}, Δy_{it} and Δq_{it}, the computed z-statistics suggest the rejection of the null hypothesis for each one of them.

In comparison, when the CIPS procedure extends to account for a vector of unobserved factors, CIPSM tests yield different results, for which the associated $\overline{CADF}$ statistics suggest the rejection of the null hypothesis of a unit root for the panel variables p_{it}, y_{it}, q_{it}. On balance the results of the panel unit root test suggest that p_{it}, y_{it}, q_{it} contain a unit root when the associated ADF regressions do not account for the cross-section dependence, and the first-difference of these variables, Δp_{it}, Δy_{it} and Δq_{it} are stationary processes.

Finally, the economy-wide real interest rate (where there are 25 time series observations) is also tested for a unit root[4]. The results are sensitive to whether or not a time trend is included in the model. In the absence of a time trend, the ADF test (with one or two lags) does not reject the null hypothesis of a unit root. When a time trend is included in the ADF regressions, the t-statistics are -4.49 for one lag and -3.27 for two lags, suggesting a trend-stationarity process for the real interest rate.

In summary, the empirical tests for the data give the evidence that the variables for the analyses contain unit roots and are first-order integrated I(1). If variables in the model specification contain unit roots I(1), a valid estimation requires the existence of one (or more) cointegrating relationships.

[4] In response to policy-induced disinflation, real interest rates appear to have trended down from the early 1990s to sometime in the 2000s but to have been more or less stable since then. The current unit root test for real interest rates did not take into account the possible structural break in the trend approximately halfway through the time period considered. It is assumed in our model that macro influences on real house price dynamics, including the impacts of real interest rates, are controlled by common factors. The possible impacts from the structural break of real interest are therefore assumed to be captured by common factors.

6.1.2 Cross-section dependence tests

In addition, the cross-section dependence tests are also conducted to verify the presence of the cross-section dependence in the data.

Table 3 reports the variable cross-section dependence and the associated CD statistics for the panel variables in the model (12). It is evident that the cross-section dependence are estimated to be significant for p_{it}, y_{it}, Δp_{it}, Δy_{it}, and q_{it}. The estimated $\bar{\rho}_{ij}$ are significantly different from 0 and the CD statistics rejects the null hypothesis of no cross-section dependence at 5% level. In the following model estimations, the possible cross-section (spatial) dependence is attempted to be accounted for by a multi-factor error structure.

6.2 Income Elasticity

The primary data analyses in section 5 reveal that the marked period 2002-2004 demonstrates the widespread rapid growth in the house prices and in the price to income ratios. As a result, an extra economic variable D_{0204} is created by using the average wage and salary income of Australia during the 2002-2004 period[5] and 0 for else. In the analysis, D_{0204} is normalized to the 2002 value. By adding this variable, the model (12) is rewritten by,

$$p_{it} = a_i + \beta_y y_{it} + \beta_q \Delta q_{i,t-1} + \beta_r r_t + \beta_d D_{0204} + \boldsymbol{\hbar}' \boldsymbol{f}_t + e_{it}, \tag{13}$$

The above model (13) is estimated with experiments to different specifications in terms of the independent variables. The vector $\boldsymbol{f_t}$ of the common shocks contains the time t cross-section averages of the observed variables $p_{it}, y_{it}, \Delta q_{i,t-1}$ and are denoted by $\bar{p}_t$, $\bar{y}_t$, and $\bar{q}_t$. The standard errors of the coefficients are estimated by the robust non-parametric estimators developed by Pesaran (2006) which are robust to both spatial and serial error correlations and continue to yield consistent estimates in the presence of them. CIPS tests where the CADF regressions are augmented with the first and second order lags respectively (reported as the $\overline{CADF(1)}$ or $\overline{CADF(2)}$ statistics) are used to test for the stationarity of the error term. The CD test is performed to test for the cross-section dependence $\hat{\bar{\rho}}$ in the error term. In particular, in consideration to the weighting issue as pointed out in the section 3, a repeated procedure is done to estimate the model in the equation (13) where the logarithm of population $q_{i,t}$ is added to control for the heterogeneity in the weights of contribution.

[5] The data is from the 1379.0.55.001 National Regional Profile Australia, 2002 to 2006.

Table 5 reports the estimates by CCE estimators. To start with, it is worth to notice that the loadings on the common factors $\bar{p}_t$ and $\bar{y}_t$ are sizable and statically significant, yet the loading on the factor $\bar{q}_t$ is not statistically significant. The economic shock D_{0204} and the interest rates do not demonstrate significant results despite showing expected signs, which may indicate that by loading on the common factors, the impacts of the short-term unobserved economic shocks as well as the state-wide unobserved factors could have been effectively captured or eliminated in the estimations which can increase the estimation accuracy.

Overviewing the results, CCEMG estimators that take into account the heterogeneities across the locations in general yield better results than the CCEP estimators, implying the heterogeneous slope coefficients for different LGAs. By CCEMG estimators, the $\overline{CADF(2)}$ statistics are sufficiently negative to reject the null hypothesis of a unit root in favour of a cointegrating relationship, and the associated estimates of the cross-section dependence $\hat{\bar{\rho}}$ are effectively zero. However, by CCEP estimators, the $\overline{CADF(2)}$ statistics can not reject the null hypothesis of a unit root at a significant level for the most regressions in experiments. As a result, the results by CCEMG estimators are preferred. Before giving weights to the LGAs, CCEMG reports the estimates of income elasticity at the range of 0.91 to 0.99 and the model that yields the estimate of 0.93(0.2) performs the best with the relative best $\bar{R}^2 = 0.95$. The model is presented as the model (3) in the Table 5. In consideration of the possible weighting issue, this model is further estimated with the LGAs weighted by the population, that is, the logarithm of population q_{it} is added to the model. In the model (5) in Table 5, it is found that the estimate of income elasticity is increased to 1.07(0.19) with the $\bar{R}^2 = 0.95$ unchanged. The increase in the estimate is likely caused by the fact that higher weights are given to LGAs with more transactions in their housing markets and with higher house prices where the income elasticities are driven to be higher. For the following analysis, 1.07 is picked up as the estimate for the income elasticity.

With the economic model being estimated, another interesting question would be the variations in the estimates over time and across locations. I then study the cross-sectional average of the residuals of the model (5) over time to look at the time variations of the estimates. Time with higher (low) residuals is characterized as outperforming (underperforming) others with respect to housing prices and the elasticity. The results are plotted in Figure 5, suggesting that even if house prices can deviate from the equilibrium relationship due to the state-wide common shocks and the temporary shocks in the economic factors, they will eventually revert. When house prices are above the equilibrium, they will tend to fall in relation to income and other economic factors and vice versa if they are falling below the equilibrium. Undoubtedly, the income elasticity varies over time due to the fluctuations in both the house prices and income. However, the income elasticity in the long run should remain to be stable, consistent with the cointegrating relationship.

To study the cross-sectional variations in the estimate of income elasticity, I look at the time-average of the residuals of the model (5) across the LGAs to study the cross-sectional deviations of the estimates. LGAs with higher (low) residuals are characterized as out-performing (underperforming) others with respect to housing prices and the elasticity. The results (unreported) show that there is no specific spatial pattern in the distribution of the population-weighted estimates. For comparison, I then also study the time-average of the residuals of the model (3) (without considering the weights). The results are reported by a GIS map in the Figure 6. The Figure 6 shows that the income elasticities across the locations demonstrate a spatial pattern, which is the highest in the Sydney (metropolitan), less significant in its surroundings and the coastal-line areas, and further diminishing as going to the inland regional areas. The findings prove the heterogeneities across the LGAs. The equilibrium positions across LGAs are not necessarily to be homogeneous but on average, there should be co-movements between the variables that lead to the state-wide cointegrating relationship.

For the estimation, Fixed-effects (FE) and mean group (MG) estimators are also used for comparisons, for which the error terms are estimated by a non-parametric variance matrix estimator that adapts the Newey and West (1987) heteroskedasticity autocorrelation consistent (HAC) procedure. The results are reported in Table 6. In comparison with the results by CCE estimators, the income elasticity estimates are relatively larger arranging from 1.53 to 1.74 which are similar to the estimates by Abelson and Chung (2005) for Australia. For the various model specification using these estimators, the effects from the interest rate and the short-term economic shock D_{0204} are significantly different from zero and presented with expected signs. However, one problem with the significance of the short-term economic shock D_{0204} is that it can increase the inaccuracy of the estimate for the income elasticity. The other problem with these estimations is that the estimates are likely to be biased due to the significant cross-section dependence in the error terms. As shown by the test results, across all the results using these estimators, $\hat{\bar{\rho}}$ is significant and the associated CD statistics strongly rejects the null hypothesis of no cross-section dependence.

6.3 Cointegration of Real House Prices and Real Income

With the estimate of income elasticity from the last section, the long run relationship between real house prices and real income to be tested for the cointegrating relationship is given by,

$$\tilde{e}_{it} = p_{it} - \hat{\beta}_y y_{it} - \hat{a}_i \tag{14}$$

where $\hat{\beta}_y = 1.07$ is the estimate of income elasticity from the model (5) in Table 5 in the last section; $\hat{a}_i$ is the estimated LGA-specific fixed effect for the ith LGA. CIPS panel unit root tests are conducted to examine the stationarity of $\tilde{e}_{it}$, for which CADF(k) regressions include the LGA specific intercepts, for different augmentation and lag orders $k = 0, 1, 2$, respectively. The CIPS statistics are computed as the simple average of the LGA-specific CADF(k) statistics. The test results reported in Table 7 suggest that the statistics for CADF(0) or CADF(1) reject the null hypothesis of a unit root in $(p_{it} - \hat{\beta}_y y_{it})$ at the 1 percent level of significance. This empirical result appears to provide some evidence for the existence of the cointegration between real house prices and real income over the observed years for the housing market in New South Wales.

Figure 7 plots the time profile showing the detrended short run dynamics of this relationship by region, computed as $\Delta(p_{i,t-1} - \hat{\beta}_y * y_{i,t-1})$ where $\hat{\beta}_y = 1.07$. It appears that the positive changes in the relationship are followed by the negative reverts to maintain the general equilibrium within the regions. Fluctuations in this relationship appear to be similar to the pattern of price changes. The next section will turn to use the error correction models to investigate the dynamics and the presence of the causality of the relationship.

6.4 Panel Error Correction Models

The empirical analyses in the previous sections 6.2 and 6.3 have confirmed the cointegrating relationship between the real house prices and the real income in the long-run. Granger (1986) and Granger (1988) imply that as long as the two variables have a common trend, the Granger causality must exist in at least one direction either unidirectional or bidirectional. Given the non-rejection of cointegration between the real house prices and the real income in our empirical analysis, the existence of the co-movements between the real house prices and the real income is confirmed for our study period, which rules out Granger non-causality and implies at least one way of Granger-causality, either unidirectional or bidirectional. Engle and Granger (1987) suggest that once a number of variables are found to be cointegrated, a corresponding error-correction representation can be constructed to detect the direction of the Granger causality through the significance of the error correction term (ECT) that is formulated from the cointegrating vector and contains the long run information. In addition, the significance of the changes in the explanatory variables and the significance of the error correction term (ECT) can also help to distinguish between the short-run and the long-run Granger causality of the variables. In order to examine the Granger causality between the real house prices and the real income and to study the dynamics in the cointegrating relationship, a vector of the error correction models (VECMs) is studied and written in equations (15-16) below,

$$\Delta p_{it} = \hbar_i + \omega \Delta y_{i,t} + \phi(p_{i,t-1} - \hat{\beta}_y y_{i,t-1}) + \varphi \Delta p_{i,t-1} + \mu_{it} \quad (15)$$

$$\Delta y_{it} = \tilde{\hbar}_i + \tilde{\omega} \Delta p_{i,t} + \tilde{\phi}(p_{i,t-1} - \hat{\beta}_y y_{i,t-1}) + \tilde{\varphi} \Delta y_{i,t-1} + \tilde{\mu}_{it} \quad (16)$$

where $(p_{i,t-1} - \hat{\beta}_y y_{i,t-1})$ is the cointegrating vector with $\hat{\beta}_y = 1.07$ and is interpreted as the departure of the current level of price from its long-run equilibrium value. $\{\hbar_i, \tilde{\hbar}_i\}$ control the regional fixed effects, $\{\hat{\omega}, \hat{\tilde{\omega}}\}$ measure the short-run impacts, $\{\hat{\phi}_i, \hat{\tilde{\phi}}\}$ measure the speed of the long-run gravitation towards the equilibrium relationship, $\{\hat{\varphi}, \hat{\tilde{\varphi}}\}$ capture the first order autocorrelation of the variables, $\{\mu_{it}, \tilde{\mu}_{it}\}$ capture the unobserved transitory shocks. Δp_{it} and Δy_{it} are first differenced terms of p_{it} and y_{it} and both of them are examined to be I(0) in section (6.1.1). In a variety of experiments for (15-16), I also include Δq_{it}, and a linear combination of the real interest rate and the time trend to the model, which forms to be I(0) as proved by the ADF test in Section 6.1.1.

Aligning with the causal hypothesis, the significance of the ECT in the equation (15) captures the unidirectional causality from real income to real house prices and the significance of the ECT in the equation (16) would capture the causality from real house prices to real income. A summary of the estimated VECMs is presented in Table 8-9.

In Table 8, CCEMG estimators yield slightly higher $\bar{R}^2$ statistics than CCEP estimators (around 0.5 to 0.3). These are likely due to the heterogeneous slopes in the estimations. It is notable that $\hat{\omega}$ are consistently estimated to be positively signed to be around 0.3 to 0.4 and $\hat{\phi}$ to be negatively signed around 0.4 to 0.5 by CCEMG and both of them are statistically significant. In addition, the real interest rate consistently and significantly exact the negative effect to the price changes[6]. In Table 9, however, it is interesting to note that even though $\hat{\tilde{\phi}}$ for the ECT and $\hat{\tilde{\omega}}$ for $\Delta p_{i,t}$ are statistically significant but the impacts are insignificant with the values approximating to 0.

Based on the analyses, as evident by the significance of $\{\hat{\phi}, \hat{\omega}\}$ in the equation (15) and the insignificance of $\{\hat{\tilde{\phi}}, \hat{\tilde{\omega}}\}$ in the equation (16), the empirical evidence is found that there is the unidirectional causal effect both 'short-term' and 'long-term' from the real income to the real house prices, which implies the within-sample Granger exogeneity of real house prices to the real income, and the Granger endogeneity of real income to the real house prices. The economic intuition underlying these results is that when real house prices are

[6]In order for comparison, MG and FE estimates are also estimated but the results are not reported since across all the experiments analyzed by the MG and FE estimators, the CD statistics again prove the strong residual cross-section dependence and the less significant estimates of $\bar{R}^2$.

subject to some either transitory or permanent shocks, the changes in the real income act to be the brunt in the adjustment of clearing the disequilibrium. The diversion from the long run equilibrium relationship would then be gradually adjusted by the convergence that drives the movement towards the long-run equilibrium.

6.5 Factor Loadings by Local Government Areas

It could be seen from the previous analysis that a cointegrating relationship between real house prices and real income is verified by adding a multi-factor structure to the econometric models where the multi-factor structure filters out the cross-section dependence in the variables. It is well-known that latent common factors would help to reduce the data required to explain a complex correlation, for example, the latent spatial patterns for panel studies. In the current analysis, the common factors are assumed to describe the economy-wide shocks that are induced by the macroeconomic conditions, policies and technologies. As a result, this would naturally lead to another interesting question - if the common factors are significant to house prices across the geographical locations given the panel study for LGAs across the state. As a matter of fact, the unobservable common factors in the current analysis are approximated by the state-wide average prices and income. I then estimate the factor loadings (the correlation coefficients) between the cointegrating vector by LGAs and the latent common factor which is formed by the cross-sectional mean of the cointegrating vector across all the LGAs in the state. The model is written as followed:

$$(p_{i,t-1} - \hat{\beta}_y * y_{i,t-1}) = \iota_i + \hat{\gamma}_i(\bar{p}_{t-1} - \hat{\beta}_y * \bar{y}_{t-1}) + \tau_{it} \quad (17)$$

$\hat{\gamma}_i$ is the factor loading for the ith LGA; $(p_{i,t-1} - \hat{\beta}_y * y_{i,t-1})$ is the cointegrating vector for the ith LGA where $\hat{\beta}_y = 1.07$ is the estimated income elasticity from the model presented in the 5th column of the Table 5. By construction, the cross-sectional average of the estimated factor loadings on the common factor $(\bar{p}_t - \hat{\beta}_y \bar{y}_t)$ is unity and the cross-sectional average of the intercepts ι_i is zero.

The factor loadings are reported by the GIS map in the Figure 8, which shows an interesting spatial pattern in the estimated loadings on the factor $(\bar{p}_t - \hat{\beta}_y \bar{y}_t)$. To start with, except the Greater Hume Shire in the Murry region and the Hunters Hill in the Sydney, the loadings on the factor $(\bar{p}_t - \hat{\beta}_y \bar{y}_t)$ are all statistically significant. Notably, all the LGAs in the Sydney, the Sydney surrounds, the coastal-line areas, and the regional areas including the Hunter region, the Mid North Coast region, the Richmond-Tweed region, the most of the South-Eastern region, and the many of the Murrumbidgee region have loadings more than 0.7. These estimates are in comparison to the loadings for the

LGAs in the inner west of the state, the most of which are less than 0.5. The common factors which could possibly be attributable to economy-wide shocks, for example, the state-wide regulations and new technologies should have been the common thread that binds all of the LGAs in the state.

As an extension, given the widespread significance of the state-wide common factors, another question can be raised on if the common shocks to a dominant region -Sydney- of the state would be propagated spatially to other regions. I then regress the cointegrating vector for each LGA on the Sydney-wide latent common factor which is created by the cross-sectional mean of the cointegrating vector across all the LGAs in Sydney. The loadings are reported by the GIS map in the Figure 9. As a contrast to the state-wide significance of the state-wide common factors, it would be very interesting to note that the loadings on the Sydney-wide common shocks demonstrate another interesting spatial pattern, more than 0.7 in the Sydney and the LGAs that are surrounding Sydney within the commuting distances, that is, the north of Sydney as far as to the Port Stephens, the South of Sydney as far as to the Eurobodalla, and the west of Sydney as far as to the Center-West region. It appears that the impacts of the common shocks to Sydney have propagated but only to a certain spatial range surrounding the Sydney.

7 Conclusions and Discussions

This paper firstly studies the theoretical relationship between real house prices and real income by using the user cost equilibrium condition and then empirically estimate the relationship by using panel data over a relatively long-term time period from 1991 to 2015 annually across 144 LGAs in New South Wales of Australia.

In the studies, the income elasticity of house prices for the state is estimated to be 1.07 by multi-factor panel data models and the cointegration analysis. The income elasticity of 0.93 reflects a relative inelastic response of the changes in the demand for housing to the income changes, even though the population-weighted income elasticity (1.07) suggests a mildly elastic response. The Granger causality of the relationship is proved to be unidirectional from real income to real house prices.

Across the state, the income elasticities are higher in the Sydney (Greater Metropolitan) and the Sydney Surrounds while gradually diminish as going to the inland regional and rural areas, suggesting that the price changes are relatively more responsive to the income changes in Sydney than anywhere else in NSW. The economic intuition is that household income in Sydney has a relatively stronger effect on the probability of being an owner and on the demand for housing services. In the rural and regional NSW, the income effects to the house prices are expected to be less significant and the household income should be less associated with the demand for housing services.

The factor loadings on the state-wide common factor are statistically significant (more than 0.7) over a majority of areas in the state in particular over the coastal areas, which possibly imply that the state-wide common factors are the common threads that bind all of the LGAs in the state. In contrast, the Sydney-wide common shocks show impacts that only propagated to a certain spatial range surrounding the Sydney. The spatial pattern for the Sydney-wide common factors is distinct, which may help to support the the central role and the existence of "ripple effect" of the Sydney house prices. However, more research is required to confirm the existence of such spatial impacts.

Acknowledgement

This research was supported by the Australian Research Council Centre of Excellence in Population Ageing Research (project number CE110001029) and Australian Laureate Fellowship FL110100247. Precious advice were received from Glenn Otto in the procedure of this research.

8 Econometric Tests

8.1 the CD Test

Suppose y_{it} is a symbol for any panel data variable requiring the test, the sample estimate of the pair-wise correlation $\hat{\rho}_{ij}$ of the ith and jth cross-section units of y_{it} is,

$$\hat{\rho}_{ij} = \hat{\rho}_{ji} = \frac{\sum_{t=1}^{T} y_{it} y_{jt}}{(\sum_{t=1}^{T} y_{it}^2)^{1/2} (\sum_{t=1}^{T} y_{jt}^2)^{1/2}} \tag{18}$$

The value of cross-section dependence is then computed as,

$$\bar{\rho}_{ij} = \frac{2}{N(N-1)} \sum_{i=1}^{N-1} \sum_{j=i+1}^{N} \hat{\rho}_{ij}. \tag{19}$$

The null hypothesis is that y_{it} is independently and identically distributed across cross-section units, written as,
$H_0 : \rho_{ij} = \rho_{ji} = 0$ for $i \neq j$,
$H_1 : \rho_{ij} = \rho_{ji} \neq 0$ for $i \neq j$,

The CD statistic proposed by Pesaran (2004) has the following form,

$$CD = \sqrt{\frac{2T}{N(N-1)}} \sum_{i=1}^{N-1} \sum_{j=i+1}^{N} \hat{\rho}_{ij} \rightarrow N(0,1), \tag{20}$$

for $N \rightarrow \infty$ and T sufficiently large.

8.2 z-statistics

This z statistic is calculated as

$$z = \frac{1}{\sqrt{N}} \sum_{i=1}^{N} \Phi^{-1}(pv_i) \tag{21}$$

where pv_i is the p-value of an Augmented Dickey Fuller (ADF) test for a unit root on the ith panel, $\Phi^{-1}(.)$ is the inverse function of the standard normal cumulative distribution. The null hypothesis is that each of the cross-section times series has a unit root, $z \sim N(0,1)$. The null hypothesis is rejected for values of $z < C_{za}$-the lower tail of the normal distribution.

8.3 CIPS and CIPSM Panel Unit Root Tests

The null hypothesis for the CIPS test is that all of the time series in the panel have a unit root and the alternative allows for some series to be stationary and some (N_1) to have unit roots where N_1/N represents the fraction of the individual processes that are stationary and tends to the fixed value Θ such that $0 < \Theta \leqslant \infty$ as $N \rightarrow \infty$. The test is performed by running a cross-section augmented Dickey Fuller test (CADF) on each of the N time series in the panel. For the ith cross-section unit the CADF test regression takes the following form;

$$\Delta y_t^i = r_0^i + r_1^i y_{t-1}^i + r_2^i \bar{y}_{t-1} + r_3^i \Delta y_{t-1}^i + r_4^i \Delta \bar{y}_{t-1} + \omega_t^i, \;\; \omega_{it} \sim i.i.d(0, \sigma^2) \tag{22}$$

where y_t^i is the data that is tested for the panel unit root, $\bar{y}_{t-1}$ and $\Delta\bar{y}_{t-1}$ are cross-section averages for the N units at time t. The CADF test regression can be generalized to include trends and additional lags of the dependent variable. Pesaran (2006) suggests that the CADF regressions are able to filter out any cross-section dependence caused by the unobserved common factors by augmenting the standard ADF regression with the cross-section averages $\bar{y}_{t-1}$ and $\Delta\bar{y}_{t-1}$. The associated CADF statistic is the t-statistic for r_1^i, denoted as $\widetilde{t}_i$. The CIPS test statistic is obtained as the cross-section average of the N values of $\widetilde{t_i}$, denoted as

$$\overline{CADF} = N^{-1} \sum_{i=1}^{N} \widetilde{t_i} \tag{23}$$

In the CIPSM test, the CIPS procedure is extended to allow for a vector of unobserved factors $\boldsymbol{f}_t$ (Pesaran, Vanessa Smith, and Yamagata 2013). The additional unobserved factors are proxied by the stationary regressors $\Delta \boldsymbol{X}_t^i$ and their cross-section average $\overline{\Delta \boldsymbol{X}_t^i}$ in the CADF test regressions. In the analysis, $\Delta \boldsymbol{X}_t^i \equiv \{\Delta q_{it}, \Delta y_{it}, \Delta p_{it}\}$.

References

Abelson Joyeux, Milunovich and Chung (2005). "Explaining House Prices in Australia: 1970–2003*". In: *Economic Record* 81.s1, S96–S103.

Abelson, Peter (1994). "House prices, costs and policies: An overview". In: *Economic Papers: A journal of applied economics and policy* 13.1, pp. 76–96.

Alexander, Carol and Michael Barrow (1994). "Seasonality and cointegration of regional house prices in the UK". In: *Urban Studies* 31.10, pp. 1667–1689.

Anselin, Luc (1988). "Lagrange multiplier test diagnostics for spatial dependence and spatial heterogeneity". In: *Geographical analysis* 20.1, pp. 1–17.

Anselin, Luc, Julie Le Gallo, and Hubert Jayet (2008). "Spatial panel econometrics". In: *The econometrics of panel data*. Springer, pp. 625–660.

Ashworth, John and Simon C Parker (1997). "Modelling regional house prices in the UK". In: *Scottish Journal of Political Economy* 44.3, pp. 225–246.

Bai, Jushan and Serena Ng (2004). "A PANIC attack on unit roots and cointegration". In: *Econometrica* 72.4, pp. 1127–1177.

Baltagi, Badi H, Seuck Heun Song, and Won Koh (2003). "Testing panel data regression models with spatial error correlation". In: *Journal of econometrics* 117.1, pp. 123–150.

Bourassa, Steven C and Patric H Hendershott (1995). "Australian capital city real house prices, 1979–1993". In: *Australian Economic Review* 28.3, pp. 16–26.

Choi, In (2001). "Unit root tests for panel data". In: *Journal of international money and Finance* 20.2, pp. 249–272.

Cook, Steven (2003). "The convergence of regional house prices in the UK". In: *Urban Studies* 40.11, pp. 2285–2294.

De Leeuw, Frank (1971). "The demand for housing: A review of cross-section evidence". In: *The Review of Economics Statistic*, pp. 1–10.

Engle, Robert F and Clive WJ Granger (1987). "Co-integration and error correction: representation, estimation, and testing". In: *Econometrica: journal of the Econometric Society*, pp. 251–276.

Fernández-Kranz, Daniel and Mark T Hon (2006). "A cross-section analysis of the income elasticity of housing demand in Spain: Is there a real estate bubble?" In: *The Journal of Real Estate Finance and Economics* 32.4, pp. 449–470.

Frees, Edward W (1995). "Assessing cross-sectional correlation in panel data". In: *Journal of econometrics* 69.2, pp. 393–414.

Gallin, Joshua (2006). "The long-run relationship between house prices and income: evidence from local housing markets". In: *Real Estate Economics*.

Granger, Clive WJ (1988). "Some recent development in a concept of causality". In: *Journal of econometrics* 39.1-2, pp. 199–211.

Granger, CW J (1986). "Developments in the study of cointegrated economic variables". In: *Oxford Bulletin of economics and statistics* 48.3, pp. 213–228.

Hadri, Kaddour (2000). "Testing for stationarity in heterogeneous panel data". In: *The Econometrics Journal* 3.2, pp. 148–161.

Harris, Richard DF and Elias Tzavalis (1999). "Inference for unit roots in dynamic panels where the time dimension is fixed". In: *Journal of econometrics* 91.2, pp. 201–226.

Hill, Robert J and Michael Scholz (2017). "Can Geospatial Data Improve House Price Indexes? A Hedonic Imputation Approach with Splines". In: *Review of Income and Wealth*.

Hill, Robert J and Iqbal A Syed (2016). "Hedonic price–rent ratios, user cost, and departures from equilibrium in the housing market". In: *Regional Science and Urban Economics* 56, pp. 60–72.

Himmelberg, Charles, Christopher Mayer, and Todd Sinai (2005). "Assessing high house prices: Bubbles, fundamentals and misperceptions". In: *Journal of Economic Perspectives* 19.4, pp. 67–92.

Holly, Sean, M Hashem Pesaran, and Takashi Yamagata (2010). "A spatio-temporal model of house prices in the USA". In: *Journal of Econometrics* 158.1, pp. 160–173.

Kapoor, Mudit, Harry H Kelejian, and Ingmar R Prucha (2007). "Panel data models with spatially correlated error components". In: *Journal of econometrics* 140.1, pp. 97–130.

Lee, Tong Hun (1968). "Housing and permanent income: Tests based on a three-year reinterview survey". In: *The Review of Economics and Statistics*, pp. 480–490.

LeSage, James P and R Kelley Pace (2010). "Spatial econometric models". In: *Handbook of applied spatial analysis*, pp. 355–376.

Leunig, Tim and Henry Overman (2008). "Spatial patterns of development and the British housing market". In: *Oxford Review of Economic Policy* 24.1, pp. 59–78.

Levin, Andrew, Chien-Fu Lin, and Chia-Shang James Chu (2002). "Unit root tests in panel data: asymptotic and finite-sample properties". In: *Journal of Econometrics* 108.1, pp. 1–24.

Liu, Xiangling and Glenn Otto (2017). "Housing supply elasticity in local government areas of Sydney". In: *Applied Economics* 0.1, pp. 1–21.

Maddala, Gangadharrao S and Shaowen Wu (1999). “A comparative study of unit root tests with panel data and a new simple test”. In: *Oxford Bulletin of Economics and statistics* 61.S1, pp. 631–652.

Maza, L and J Pages (2007). “Analysis of house prices in Spain”. In: *Working Paper No. 0307*. Bank of Spain Spain.

Meen, Geoffrey (1999). “Regional house prices and the ripple effect: a new interpretation”. In: *Housing studies* 14.6, pp. 733–753.

Moon, Hyungsik Roger and Benoit Perron (2004). “Testing for a unit root in panels with dynamic factors”. In: *Journal of econometrics* 122.1, pp. 81–126.

Newey, Whitney K and Kenneth D West (1987). “Hypothesis testing with efficient method of moments estimation”. In: *International Economic Review* 28.3, pp. 777–787.

Pesaran, M Hashem (2004). *General diagnostic tests for cross section dependence in panels*. CESifo working paper series.

– (2006). “Estimation and inference in large heterogeneous panels with a multifactor error structure”. In: *Econometrica* 74.4, pp. 967–1012.

– (2007). “A simple panel unit root test in the presence of cross-section dependence”. In: *Journal of Applied Econometrics* 22.2, pp. 265–312.

Pesaran, M Hashem and Elisa Tosetti (2011). “Large panels with common factors and spatial correlation”. In: *Journal of Econometrics* 161.2, pp. 182–202.

Pesaran, M Hashem, L Vanessa Smith, and Takashi Yamagata (2013). “Panel unit root tests in the presence of a multifactor error structure”. In: *Journal of Econometrics* 175.2, pp. 94–115.

Poterba, James M (1984). “Tax subsidies to owner-occupied housing: an asset-market approach”. In: *The quarterly journal of economics* 99.4, pp. 729–752.

Sarafidis, Vasilis, Takashi Yamagata, and Donald Robertson (2009). “A test of cross section dependence for a linear dynamic panel model with regressors”. In: *Journal of econometrics* 148.2, pp. 149–161.

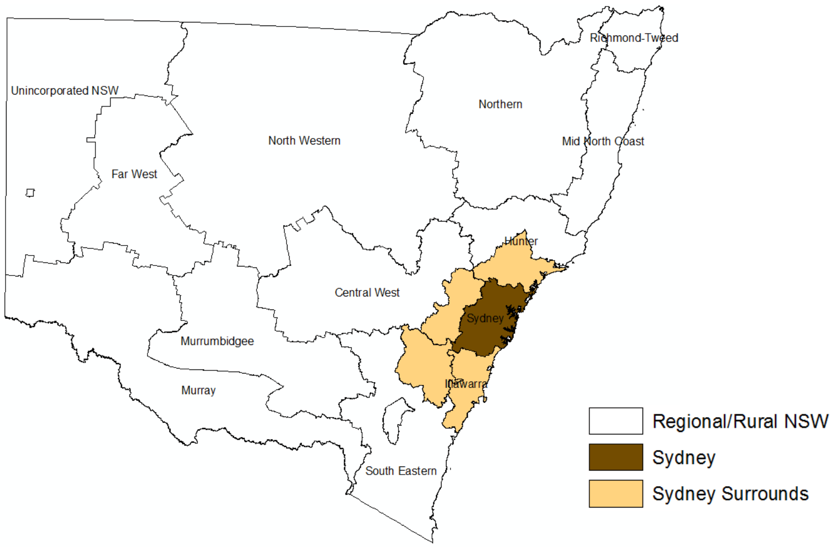
Slater, JR (1986). “Income, Location and Housing in Greater London”. In: *Urban Studies* 23.4, pp. 333–341.

Wilkinson, RK (1973). “The income elasticity of demand for housing”. In: *Oxford Economic Papers* 25.3, pp. 361–377.

Table 1: Variable Symbols and their Explanations

P_{it}	Real house prices
Y_{it}	Real income per taxpayer
p_{it}	Logarithm of real house prices
y_{it}	Logarithm of real income per taxpayer
q_{it}	Logarithm of population
r_t	Real interest rate
Δp_{it}	The first difference of Logarithm of real house prices
Δy_{it}	The first difference of Logarithm of real income per taxpayer
Δq_{it}	The first difference of Logarithm of population
D_{0204}	The transitory demand shock during 2002 to 2004, proxied by an estimate for the permanent income.

Figure 1: Regions of New South Wales, Australia and the three LGA-groupings



Notes. The GIS map is generated by overlapping the Regions of New South Wales, Australia (NSW) with the three LGA groupings used in the analysis. The two of the three LGA groupings are "Sydney" and "Sydney Surrounds", coloured by the dark yellow and the light yellow respectively. The uncoloured area (white) is the rest of NSW, grouped as "Regional and Rural NSW". The LGAs each area contains are listed in Table 2. The GIS Map for the LGAs of New South Wales, Australia can be found on the website: http://www.screen.nsw.gov.au/page/maps/nsw-state-council-map.

Table 2: The List of Regions and Local Government Areas in New South Wales, Australia

Sydney	Tumut Shire	Hastings
Ashfield	Upper Lachlan*	Kempsey
Auburn	Yass Valley	Nambucca
Bankstown	Young	**Richmond-Tweed**
Baulkham Hills	**Central West**	Ballina
Blacktown	Bathurst Regional	Byron
Blue Mountains	Bland	Kyogle
Botany Bay	Blayney	Lismore
Burwood	Cabonne	Richmond Valley
Camden	Cowra	Tweed
Campbelltown	Forbes	**Murrumbidgee**
Canada Bay	Lachlan	Carrathool
Canterbury	Lithgow*	Coolamon
Fairfield	**Mid-Western Regional**	Cootamundra
Gosford	Oberon*	Griffith
Hawkesbury	Orange	Gundagai
Holroyd	Parkes	Hay
Hornsby	Weddin	Junee
Hunter's Hill	**Illawarra**	Leeton
Hurstville	Kiama*	Lockhart
Kogarah	Shellharbour*	Murrumbidgee
Ku-ring-gai	Shoalhaven*	Narrandera
Lane Cove	Wingecarribee*	Temora
Leichhardt	Wollongong*	Wagga Wagga
Liverpool	**Hunter**	**North Western**
Manly	Cessnock*	Bogan
Marrickville	Dungog	Bourke
Mosman	Gloucester	Brewarrina
North Sydney	Great Lakes	Cobar
Parramatta	Lake Macquarie*	Coonamble
Penrith	Maitland*	Dubbo
Pittwater	Muswellbrook	Gilgandra
Randwick	Newcastle*	Narromine
Rockdale	Port Stephens*	Walgett
Ryde	Singleton*	Warren
Strathfield	Upper Hunter Shire	Warrumbungle Shire
Sutherland Shire	**Murray**	Wellington
Sydney	Albury	**Northern**
Warringah	Balranald	Armidale Dumaresq
Waverley	Berrigan	Glen Innes Severn
Willoughby	Conargo	Gunnedah
Wollondilly	Corowa Shire	Guyra
Woollahra	Deniliquin	Gwydir
Wyong	Greater Hume Shire	Inverell
South Eastern	Jerilderie	Liverpool Plains
Bega Valley	Murray	Moree Plains
Bombala	Tumbarumba	Narrabri
Boorowa	Urana	Tamworth Regional
Cooma-Monaro	Wakool	Tenterfield
Eurobodalla	Wentworth	Uralla
Goulburn Mulwaree*	**Mid North Coast**	Walcha
Harden	Bellingen	**Far West**
Palerang	Clarence Valley	Broken Hill
Queanbeyan	Coffs Harbour	Central Darling
Snowy River	Greater Taree	

Notes: LGAs are listed in the alphabetical order within each region. The "Sydney" in the study correspond to the region of "Sydney" in the table, "Sydney Surrounds" is comprised by the LGAs with the "*" mark. The "Regional/Rural NSW" includes all the other LGAs in the NSW, that is, all the LGAs except the LGAs in "Sydney" and the * marked LGAs. Referring to Figure 1 for the position of the three areas for analyses. Regional LGAs including Balranald, Brewarrina, Central Darling, Conargo, Jerilderie, Murrumbidgee, Urana, Gwydir are dropped from the analysis (and this list) due to their missing values of house prices for the majority of years between 1991 to 2015.

Table 3: Variable Cross-section Dependence Testing Results

Variable	CD-Statistics	$\bar{\hat{\rho}}_{ij}$	Variable	CD-Statistics	$\bar{\hat{\rho}}_{ij}$
p_{it}	370.81	0.78	Δp_{it}	139.60	0.30
y_{it}	453.33	0.95	Δy_{it}	213.19	0.46
q_{it}	69.75	0.15	Δq_{it}	27.78	0.06

Notes. The cross-section dependence $\bar{\hat{\rho}}_{ij}$ is defined as the average of the pair-wise correlation, $\hat{\rho}_{ij}$, between the cross-section units of the variable. Test statistics for cross-section dependence is $CD = \sqrt{\frac{2T}{N(N-1)}} \sum_{i=1}^{N-1} \sum_{j=i+1}^{N} \hat{\rho}_{ij} \xrightarrow{d} N(0,1)$.

Table 4: Panel Unit Root Testing Results

	Z-test (1)	Z-test (2)	CIPS(1)	CIPS(2)	CIPSM(1)	CIPSM(2)
With an intercept						
p_{it}	4.82	6.91	-1.82	-1.51	-2.32	-1.56
y_{it}	14.64	17.49	-2.06	-1.48	-2.32	-1.85
q_{it}	4.95	5.62	-1.69	-1.07	-2.24	-1.91
Δp_{it}	-22.15	-12.74	-3.35	-2.31	-3.45	-2.20
Δy_{it}	-30.22	-20.65	-3.14	-2.25	-3.10	-2.06
Δq_{it}	-15.07	-10.37	-2.39	-1.83	-3.11	-2.61
With an intercept and a linear trend						
p_{it}	3.76	6.39	-2.28	-1.63	-2.39	-1.77
y_{it}	-3.79	1.37	-2.04	-1.53	-1.93	-1.54
q_{it}	4.62	4.30	-1.69	-1.64	-2.82	-2.17
Δp_{it}	-14.50	-5.33	-3.63	-2.62	-3.88	-2.80
Δy_{it}	-23.87	-15.88	-3.57	-2.74	-3.61	-2.80
Δq_{it}	-11.90	-8.94	-3.00	-2.31	-3.54	-2.57

Notes. Panel unit root tests for each variable are conducted for the time lag order k of ADF regressions. The reported statistics of z-test are the inverse normal z-statistics computed by (21). The z-test is the lower-tailed test. Critical values are -2.326, -1.645, -1.282 for 1 %, 5% and 10% significance respectively. CIPS(k) statistics are computed as the simple average of the individual-specific CADF(k) statistics using (22), for augmentations with the different lag order k in the CADF regressions. Critical values for CIPS or CPISM tests are -2.14, -2.04 and -1.99 for 1 %, 5% and 10% significance respectively in the case of an intercept only, and -2.65, -2.55 and -2.49 for 1%, 5% and 10% significance respectively in the case of an intercept and a linear trend.

Table 5: Estimation Results: Income Elasticity of Real House Price: 1991-2015

p_{it}	CCEMG (1)	CCEP (2)	CCEMG (3)	CCEP (4)	CCEMG (5)*	CCEP (6)
$\hat{\beta}_y$	0.99	0.74	0.93	0.69	1.07	0.78
	(0.21)***	(0.22)***	(0.20)***	(0.20)***	(0.19)***	(0.22)**
$\hat{\beta}_r$	-0.00	-0.00	-0.00	-0.00	0.00	0.00
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
$\hat{\beta}_d$	0.00	0.00	0.00	0.00	0.00	0.00
	(0.02)	(0.02)	(0.02)	(0.02)	(0.01)	(0.02)
$\hat{\beta}_q$			0.17	-0.11	0.22	-0.08
			(0.30)	(0.19)	(0.33)	(0.22)
$\hat{q}_w$					0.22	0.34
					(0.28)	(0.24)
$\hbar_{\bar{q}}$			0.88	-0.11	1.28	0.08
			(1.22)	(1.35)	(1.1)	(1.28)
$\hbar_{\bar{p}}$	0.97	0.99	0.97	1.00	0.98	1.00
	(0.04)***	(0.03)***	(0.04)***	(0.03)***	(0.04)***	(0.03)***
$\hbar_{\bar{y}}$	-0.99	-0.74	-0.92	-0.69	-1.02	-0.78
	(0.21)***	(0.06)***	(0.20)***	(0.06)***	(0.23)***	(0.19)***
$\overline{CADF(2)}$	-2.90	-1.20	-2.98	-1.96	-3.83	-2.45
CD	2.66	0.37	2.38	-0.04	5.68	0.87
$\hat{\bar{\rho}}$	0.01	0.01	0.00	-0.01	0.01	0.00
$\bar{R}^2$	0.87	0.80	0.95	0.92	0.95	0.92

Notes. The general estimation model is $p_{it} = a_i + \beta_y y_{it} + \beta_q \Delta q_{i,t-1} + \beta_r r_t + \beta_d D_{0204} + \hbar_{\bar{p}} \bar{p}_t + \hbar_{\bar{y}} \bar{y}_t + \hbar_{\bar{q}} \bar{\Delta} q_{t-1} + e_{it}$, where $\bar{p}_t, \bar{y}_t, \bar{\Delta} q_{t-1}$ are the time t cross-section averages of p_{it}, y_{it} and $\Delta q_{i,t-1}$. The model (5-6) add the logarithm of population $q_{i,t}$ and $\hat{q}_w$ is the estimate for the coefficient of $q_{i,t}$. Numbers in the parenthesis are robust standard errors. $\overline{CADF(2)}$ are the CIPS(2) statistics. Critical values for CIPS(2) tests are -2.14, -2.04 and -1.99 for 1%, 5% and 10% significance respectively in the case of an intercept only. CD is the statistics testing for the significance of cross-section dependence computed by (20). $\hat{\bar{\rho}}$ is the estimate for the average of the pair-wise correlations of the cross-section units computed by (19). The superscript *,**,*** signifies the test is significant at the 10, 5, 1 per cent level. The $\star$ indicates $\hat{\beta}_y$ of the model is used for the following analysis including cointegration, ECMs and factor loadings.

Table 6: Estimation Results: Income Elasticity of Real House Price: 1991-2015

p_{it}	MG	FE	MG	FE	MG	FE
$\hat{\beta}_y$	1.74	1.53	1.72	1.54	0.98	1.43
	(-0.05)***	(0.04)***	(0.06)***	(-0.05)***	(0.15)***	(0.05)***
$\hat{\beta}_r$	-0.01	-0.01	-0.01	-0.01	-0.02	-0.01
	(0.002)***	(0.002)***	(0.002)***	(0.002)***	(0.00)***	(0.00)***
$\hat{\beta}_q$			-2.35	-0.20	-1.29	-0.24
			(-0.59)***	(-0.24)***	(0.54)**	(0.23)
$\hat{\beta}_d$	0.07	0.06	0.06	0.06	0.03	0.06
	(-0.02)***	(-0.01)***	(-0.02)***	(-0.01)***	(0.01)**	(0.01)***
$\hat{q}_w$					0.70	0.47
					(0.43)	(0.07)***
$\overline{CADF(2)}$	-2.60	-1.22	-2.72	-1.39	-3.33	-1.96
CD	265.32	195.12	225.52	201.70	157.73	207.97
$\hat{\bar{\rho}}$	0.52	0.38	-0.46	0.41	0.32	0.43
$\bar{R}^2$	0.66	0.57	0.85	0.80	0.88	0.80

Notes. The general estimation model is $p_{it} = a_i + \beta_y y_{it} + \beta_q \Delta q_{i,t-1} + \beta_r r_t + \beta_d D_{0204} + e_{it}$. In an alternative regression, the logarithm of population $q_{i,t}$ is added to the model and $\hat{q}_w$ is the estimate for the coefficient of $q_{i,t}$. Numbers in the parenthesis are robust standard errors. CD is the statistics testing for the significance of cross-section dependence computed by (20). $\hat{\bar{\rho}}$ is the estimate for the average of the pair-wise correlations of the cross-section units computed by (19). The superscript *,**,*** signifies the test is significant at the 10, 5, 1 per cent level.

Table 7: Cointegration Testing Results

CIPS			CIPSM		
CADF(0)	CADF(1)	CADF(2)	CADF(0)	CADF(1)	CADF(2)
With Intercept and Trend					
-2.74	-2.53	-1.64	-2.49	-2.12	-1.12
With Intercept					
-2.17	-1.97	-1.40	-2.49	-2.12	-1.24

Notes. The cointegration test is conducted for the error term of the relationship between real house prices and real income, computed by $\tilde{e}_{it} = p_{it} - \hat{\beta}_y y_{it} - \hat{a}_i$; See more details in notes to Table 4.

Table 8: Estimation Results: Panel Error Correction Models

Δp_{it}	CCEMG	CCEP	CCEMG	CCEP	CCEMG	CCEP
$\hat{\phi}$	-0.43	-0.39	-0.59	-0.39	-0.60	-0.39
	(0.03)***	(0.02)***	(0.03)***	(0.03)***	(0.03)***	(0.04)***
$\hat{\omega}$	0.69	0.12	0.75	0.12	0.74	0.12
	(0.11)***	(0.16)	(0.12)***	(0.16)	(0.12)***	(0.16)
$\hat{\varphi}$	0.14	-0.09	0.11	-0.09	0.11	-0.09
	(0.03)***	(0.05)	(0.02)***	(0.04)*	(0.02)***	(0.04)**
θ_r			-0.001	0.001	-0.001	-0.001
			(0.002)	(0.002)	(0.002)	(0.003)
θ_d			0.00	0.00	0.00	0.00
			(0.00)	(0.00)	(0.00)	(0.00)
θ_p					0.59	0.20
					(0.24)**	(0.23)
CD:	1.04	1.82	-0.77	1.82	-0.17	1.94
$\hat{\bar{\rho}}$	0.002	0.003	-0.002	0.003	0.003	-0.003
$\bar{R}^2$	0.42	0.37	0.46	0.37	0.47	0.38

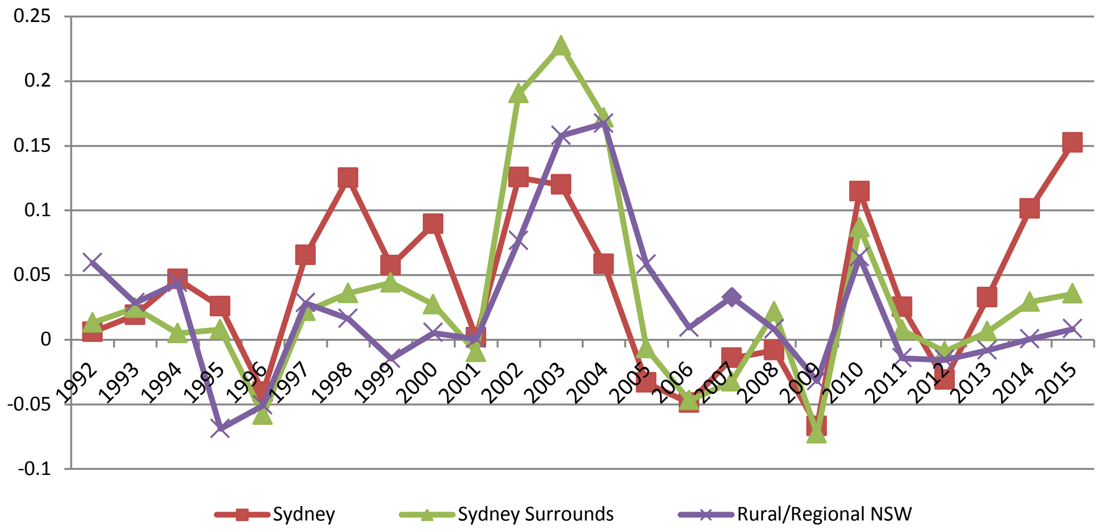
Notes. The general estimation model is $\Delta p_{it} = a_i + \phi(p_{i,t-1} - \hat{\beta} y_{i,t-1}) + \omega \Delta y_{i,t} + \varphi \Delta p_{i,t-1} + \theta_r r_t + \theta_d t + \theta_p \Delta q_{it} + \mu_{it}$. CD is the computed statistics testing for the significance of cross-section dependence computed by (20). $\hat{\bar{\rho}}$ is the estimate for the average of the pair-wise correlations of the cross-section units computed by (19).

Table 9: Estimation Results: Panel Error Correction Models

Δy_{it}	CCEMG	CCEP	CCEMG	CCEP	CCEMG	CCEP
$\tilde{\phi}$	0.00	0.00	0.01	0.00	0.01	0.00
	(0.01)	(0.01)	(0.00)*	(0.01)	(0.00)***	(0.01)
$\tilde{\omega}$	0.02	0.01	0.03	0.02	0.03	0.01
	(0.00)***	(0.01)	(0.01)***	(0.01)	(0.00)***	(0.01)
$\tilde{\varphi}$	-0.03	-0.01	-0.08	-0.01	-0.10	-0.01
	(0.02)	(0.03)	(0.02)***	(0.04)	(0.02)***	(0.04)
$\tilde{\theta}_r$			-0.00	-0.00	-0.00	-0.00
			(0.00)	(0.00)	(0.00)	(0.00)
$\tilde{\theta}_d$			-0.00	0.00	-0.00	0.00
			(0.00)	(0.00)	(0.00)	(0.00)
$\tilde{\theta}_p$					-0.08	0.01
					(-0.06)	(0.05)
CD:	14.22	10.90	13.76	10.49	14.07	10.5
$\hat{\bar{\rho}}$	0.03	0.02	0.03	0.02	0.03	0.02
$\bar{R}^2$	0.35	0.32	0.43	0.32	0.44	0.32

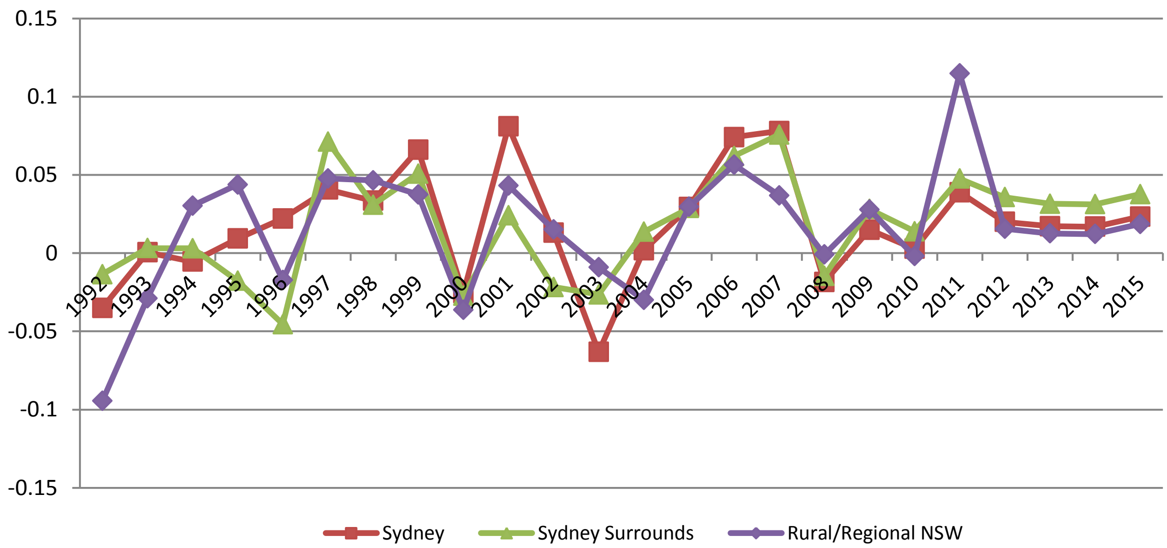
Notes. The general estimation model is $\Delta y_{it} = \tilde{a}_i + \tilde{\phi}(p_{i,t-1} - \hat{\beta} y_{i,t-1}) + \tilde{\omega} \Delta p_{i,t} + \tilde{\varphi} \Delta y_{i,t-1} + \tilde{\theta}_r r_t + \tilde{\theta}_d t + \tilde{\theta}_p \Delta q_{it} + \tilde{\mu}_{it}$. CD is the computed statistics testing for the significance of cross-section dependence computed by (20). $\hat{\bar{\rho}}$ is the estimate for the average of the pair-wise correlations of the cross-section units computed by (19).

Figure 2: Time Profiles for the Mean of Log Real House Price Changes



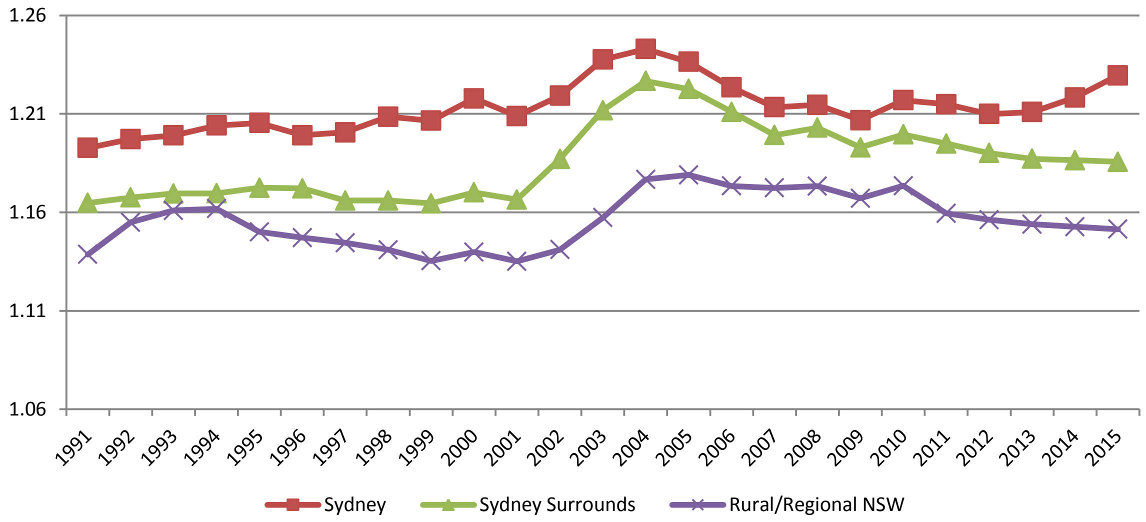
Notes. The plots are made from the average of the changes in the log real house prices, $p_{it} - p_{i,t-1}$, across LGAs within each area. See details of LGAs within each area in Table 2 and the position of each area in Figure 1.

Figure 3: Time Profiles for the Mean of Log Real Income Changes



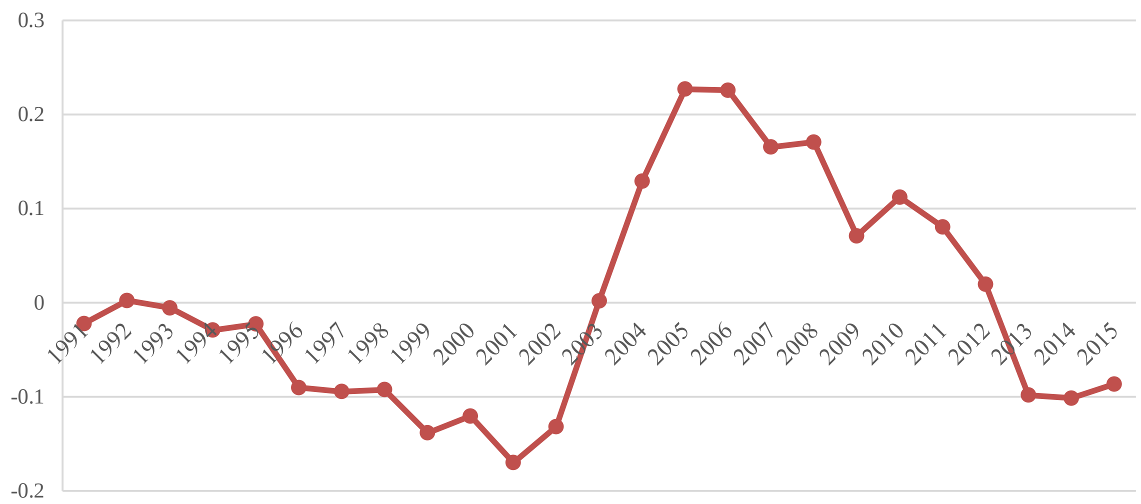
Notes. The plots are made from the average of the changes in the log real income, $y_{it} - y_{i,t-1}$, across LGAs within each area. See details of LGAs within each area in Table 2 and the position of each area in Figure 1.

Figure 4: Time Profiles for the Mean of the Ratio of Log Real House Price to Log Real Income



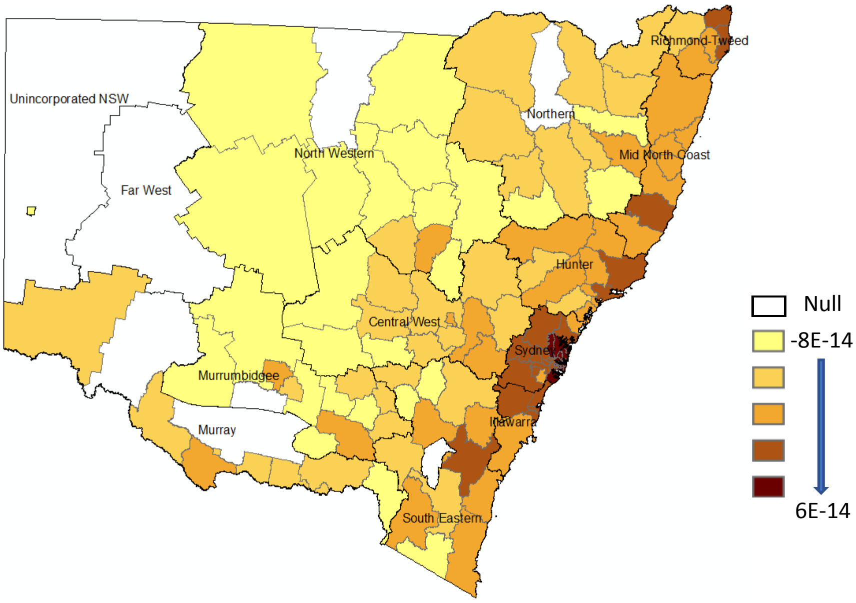
Notes. The plots are made from the average for the ratio of log real house price to log real income, p_{it}/y_{it}, across LGAs within each area. See details of LGAs within each area in Table 2 and the position of each area in Figure 1.

Figure 5: The Time Variations in the Income Elasticity



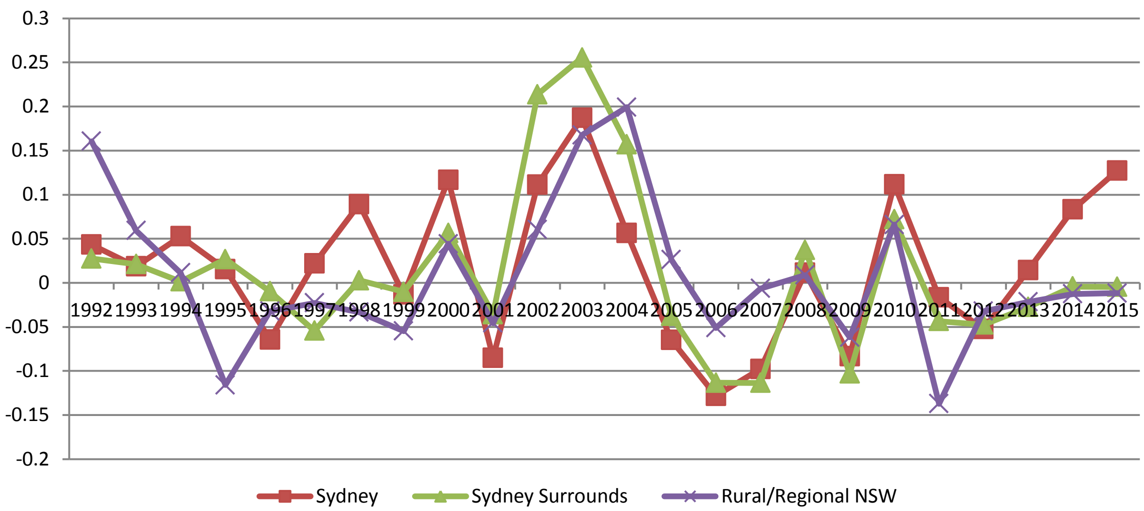
Notes. The plot is the cross-sectional mean of the residuals of the model (5) in Table 5.

Figure 6: The Cross-section Variations in the Income Elasticity



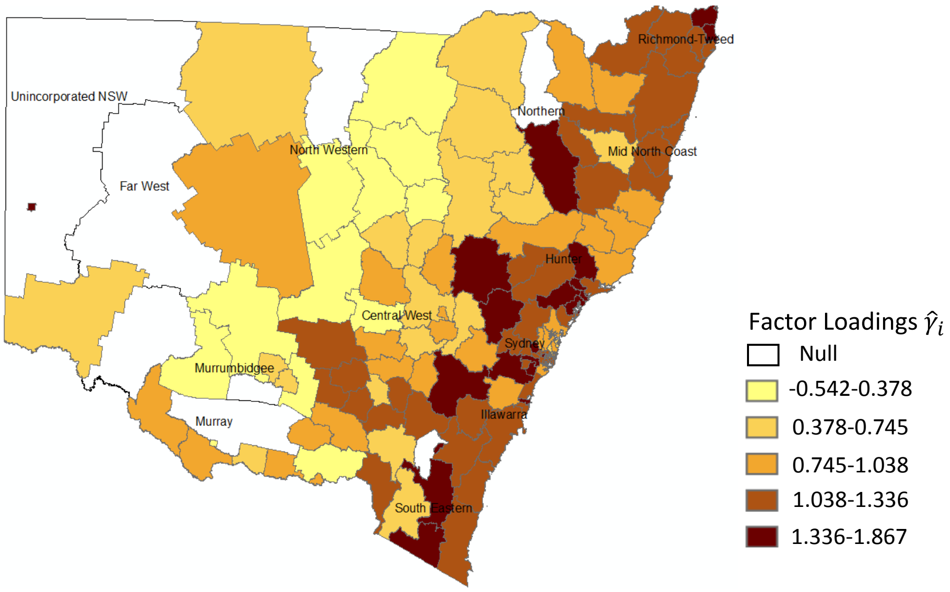
Notes: The time average of the residuals for each LGA from the model (3) in Table 5. "Null" indicates the LGA is not analysed due to data missing. See the list of LGAs in Table 2 and the GIS Map of the LGAs in New South Wales, Australia on http://www.screen.nsw.gov.au/page/maps/nsw-state-council-map.

Figure 7: Regional Time Profiles for the Changes in the Weighted Ratio of Log Real House Price to Log Real Income



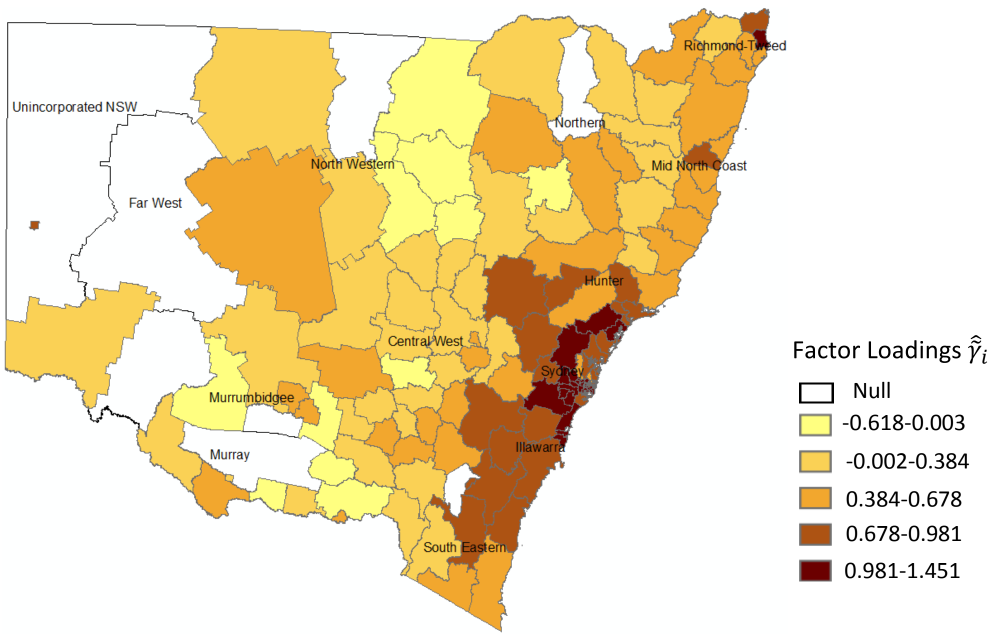
Notes. The plots are made from the average of the changes in the weighted ratio of log real house price to log real income, $\Delta(p_{it} - \hat{\beta} y_{it})$ across LGAs within each area. $(p_{it} - \hat{\beta} y_{it})$ is the cointegrating vector of p_{it} and y_{it} and $\hat{\beta} = 1.07$. See details of LGAs within each area in Table 2 and the position of each area in Figure 1.

Figure 8: The Cross-sectional Distribution of Factor Loadings on the State-wide common factor



Notes. The "Null" indicates the LGA is not analysed due to missing data. See the list of LGAs on Table 2 and the NSW state LGAs Map on http://www.screen.nsw.gov.au/page/maps/nsw-state-council-map.

Figure 9: The Cross-sectional Distribution of Factor Loadings on the Sydney-wide common factor



Notes. See notes to Figure 8.