THE UNIVERSITY OF WESTERN AUSTRALIA
Achieving International Excellence

ECONOMICS

MEASURING WESTERN AUSTRALIAN HOUSE PRICES: METHODS AND IMPLICATIONS

by

Callum Jones
Business School
The University of Western Australia

DISCUSSION PAPER 09.10

Measuring Western Australian House Prices: Methods and Implications

Callum Jones

jonescallum@gmail.com

Abstract

This paper investigates models of house price measurement with a particular focus on the hybrid hedonic repeat-sales model. It examines different ways to model house price changes and outlines a method of estimating the hybrid measure. After describing the models, the paper applies Western Australian house sales data to the outlined models and draws some conclusions as to the relative attractiveness of the hybrid measure.

Generous and helpful financial support through the Brian Gray Scholarship Program organised by the Australian Prudential Regulation Authority is warmly acknowledged, as is a C. A. Vargovic Bursary from the UWA Business School. I thank the selection committees for both scholarships. I thank numerous others for their help, including, but not limited to, Professor Ken Clements of the UWA Business School, Dr Katrina Ellis of the Australian Prudential Regulation Authority, Mr Stewart Darby of the Real Estate Institute of Western Australia, and Dr James Fogarty and Giri Parameswaran for their helpful comments. Any mistakes are mine.

Contents

1 Introduction **1**

2 The Need for Accurate Price Measures **1**

3 Models of House Prices **2**

4 Hybrid Hedonic Repeat-Sales Approach **4**

4.1 Estimation Procedure 5

5 Western Australian House Prices **6**

5.1 Models Constructed 8

5.2 Estimated Price Indexes 9

5.3 Price Index Evaluation 14

6 Conclusion **17**

References **18**

1 Introduction

This article investigates models of house price measurement with a particular focus on the advantages of the hybrid hedonic repeat-sales model ('hybrid method'). Accurate house price measures are important for policy-makers to gauge the health of this important asset class, and for consumers who need to appropriately assess their level of wealth and the risk thereof associated with purchasing a house. In the context of strong house price appreciation in Australia since 2000, these are important motivating considerations.

The article builds on recent work by Hansen (2006) (Hansen (forthcoming, 2009)) by focussing on the hybrid method. It is organised as follows: in the next section, some motivating considerations for accurate house price measures are considered. In Section 3, a brief review of well-developed house price measures is undertaken. The hybrid method is examined in Section 4. Section 5 applies these measures and Section 6 concludes.

2 The Need for Accurate Price Measures

The causal nexus of events in the United States sub-prime housing market and the resulting 'global financial crisis' illustrates the interdependence between the housing market and the wider economy. In addition to its financial market connections, housing is the largest source of wealth for individual households, with official statistics indicating dwellings constitute about 60 percent of total household wealth for the December 2007 quarter (Reserve Bank of Australia, 2008). As previous Reserve Bank of Australia Governor Ian MacFarlane (2004) says: 'housing is the biggest asset in the country ... it is an extremely important asset class for most people'.

As a large source of wealth, the macroeconomic implications of movements in the price of houses need consideration. We might hypothesise that rising house prices relax homeowners' credit constraints, allowing them to access the rising equity in the homes to finance loans for consumption; as The Economist (2005) writes in the context of the early-2000s global housing boom: 'owners ... have been using their home like an ATM machine to extract cash' (The Economist, 2005, p. 16).

A second conduit between house prices and consumption, which states that house price movements have real wealth effects, is controversial. Campbell and Cocco (2007) cite Sinai and Souleles (2005) to show how house price fluctuations do not change homeowners' real wealth since housing is a consumption good and the change in the implicit rental cost for homeowners who anticipate living in their homes for a long period of time is a perfect hedge for fluctuations in house prices.

In any event, numerous empirical studies have been conducted to investigate the relationship. In an early study Bhatia (1987) finds empirical evidence for a positive relationship between house prices and consumption. Engelhardt (1994), Engelhardt (1996) and Sheiner (1995) reach similar conclusions. More recently, Case, Quigley

Table 1: Median House Prices and Annual Growth Rates, December 2007

Data Provider	Sydney	Melb.	Bris.	Perth	Adel.	Canb.	Hobart	Darwin
A. Level Estimates								
Residex	584,500	480,000	432,500	502,000	353,000	452,500	344,500	398,500
Real Estate Institute	536,022	420,691	375,454	460,501	315,084	418,718	309,764	393,854
Adviser Edge	572,685	436,950	407,890	514,517	363,875	486,040	258,400	421,050
APM	553,357	463,488	425,368	508,776	400,659	506,570	280,853	443,917
RP-Data / Rismark	576,458	440,390	451,292	506,179	396,329	478,691	–	393,737
Lowest	536,022	420,691	375,454	460,501	315,084	418,718	258,400	393,854
Highest	584,500	480,000	451,292	514,517	400,659	506,570	344,500	443,917
Level Difference	48,478	59,309	75,838	54,016	85,575	87,852	86,100	50,063
% Difference	9%	14%	20%	12%	27%	21%	33%	13%
B. Percentage Changes (%)								
Residex	9.01	23.54	18.96	2.15	17.84	13.62	11.20	13.71
ABS	8.00	18.10	21.60	1.10	20.20	14.30	11.10	11.10
Real Estate Institute	3.40	13.10	16.20	3.40	10.70	8.10	9.30	3.90
Adviser Edge	9.50	22.10	17.90	3.90	12.70	16.00	9.50	4.70
APM	4.80	25.20	20.10	1.70	20.00	14.60	11.30	5.30
RP-Data / Rismark	5.92	18.98	21.40	-1.16	26.09	9.47	–	13.23
Lowest	3.4	13.1	16.2	-1.2	10.7	8.1	9.3	3.9
Highest	9.5	25.2	21.6	3.9	26.1	16.0	11.3	13.7
Level Difference	6.1	12.1	5.4	5.1	15.4	7.9	2.0	9.8

Source: Matusik (2008)

and Shiller (2001) find that the effect of house prices on consumption is larger than the effect of stock prices on consumption, a notable result as the three authors are important housing economics authorities. In an Australian context, Dvornak and Kohler (2003) derive support for Case et al.'s (2001) conclusions, finding a marginal propensity to consume out of housing of about 0.03.

Despite the importance of the housing market to the wider economy, current published measures of house price movements display considerable dispersion. Michael Matusik (2008) collates the variation in median house price estimates–his results are presented here in Table 1. As indicated, the dispersion is present Australia-wide. As a result, Australian citizens exposed to varying presented house price movements may base their purchasing or consumption decisions on inaccurate or incorrect information. Further, public officials may make important fiscal or monetary policy decisions which are sensitive to movements in house prices. This paper looks at more accurate measures of house prices, in an attempt to provide greater clarity in calculating estimates of changes in house prices.

3 Models of House Prices

This section outlines various models of house prices, and discusses their advantages and disadvantages. The measure most prominently reported in the press is the median house price. This has two main problems: sample compositional issues and quality bias. Under the sample compositional problem, the median can fluctuate because of changes in the make-up of the house sales sample independent of any underlying change in the demand and supply conditions. For example, if the composition of cheaper houses in low socio-economic areas increases across time, the median may shift downwards, all else equal. It is estimated this effect results in wide fluctuations in the reported median (Prasad and Richards, 2006). Indeed, Matusik's (2008) research indicates estimated median house prices do display considerable variation (Table 1). The second problem of the median measure is quality bias. Hansen (2006) reports that approximately 2% of GDP is comprised of dwelling investments, indicating quality changes can impact house price measures that do not adjust for housing quality.

The mix-adjusted measure attempts to correct for compositional problems by aggregating weighted measures of central tendency over sub-groups containing a homogenous set of houses. That is, groups $s = \{1, \ldots, S\}$ of similar houses are aggregated, a central measure P_s of each group is formed, and the groups are weighted by w_s and aggregated to form P_{MA}:

$$P_{\text{MA}} = \prod_{s=1}^{S} P_s^{w_s}.$$

Prasad and Richards (2006) show that equally weighted mix-adjusted measures differ little from a weighted index, which simplifies the analysis. The Australian Bureau of Statistics construct a mix-adjusted measure with sub-groups based on the SEIFA index of a house.

One regression method to correct for quality and compositional issues is the hedonic model. Under this approach, house prices are conceptualised as an aggregation of the prices of individual house characteristics. The time-dummy hedonic regression:

$$p_h^t = \alpha + \sum_{k=1}^{K} \beta_k x_{k,h}^t + \sum_{t=1}^{T} \gamma^t D_h^t + \varepsilon_h^t, \tag{1}$$

therefore regresses individual $h = 1, \ldots, H$ house prices p_h^t over time t against its observable characteristics $x_{k,h}^t$ and a time-dependent dummy variable D_h^t, where $D_h^t = 1$ in the time period when house h sells, and 0 otherwise. The dummy variable should capture the 'shift' in the regression line resulting purely from the passage of time (de Haan, 2007).

The parameter γ^t is the estimated change in the price of an average house, with its interpretation depending on how house prices p_h^t are included in the model. If log prices are used as the independent variable, γ^t is the approximate percentage change in house prices; if actual prices are used, γ^t gives the level change in prices. Note, a weighted hedonic regression is not used. If it were, the interpretation would be changes in the value of a representative portfolio of houses.

The repeat-sales approach to modelling house prices uses houses which sell at least twice during the sample period (define these as houses $j = 1, \ldots, J$). If that house's characteristics remain unchanged, the hedonic regression (1) for the first sale of the house can be subtracted from the hedonic specification of the second sale–the characteristics drop out from the regression and the change in prices is regressed against a series of time-dependent variables:

$$p_j^\tau - p_j^t = \sum_{\tau=1}^{T} \gamma^\tau \Delta_j^\tau + \varepsilon_j^\tau - \varepsilon_j^t, \tag{2}$$

where Δ_j^τ is a time-dependent variable equal to 1 for the second of the pair of sales for house j, -1 for the first sale of the pair and 0 otherwise. The price change, measured by γ^τ then gives the quality-adjusted house price change.

Clearly, the most attractive advantage of the repeat-sales approach is that it does not require hedonic specification of characteristics and can be implemented using much less data than a fully specified hedonic model. Equally clearly is the drawback of such an approach; in conducting regression (2) the researcher is assuming characteristics $x_{k,j}^t$ are unchanged from t to τ which is a strong assumption to make in the absence of detailed data. A second major disadvantage of the repeat-sales approach is that it throws away a high proportion of available house price information as only a fraction of houses in any sample are repeat-sales observations (in the dataset used in this analysis, about one-third of house sale observations are repeat-sales).

Generally then we see a trade-off between the two regression approaches, offsetting the demand of hedonic specification with the strength of repeat-sales observations represented in the house sales sample. The hybrid hedonic repeat-sales approach is an attractive complementary approach which, as its name suggests, combines the best features of the hedonic and repeat-sales approaches.

4 Hybrid Hedonic Repeat-Sales Approach

The hybrid hedonic repeat-sales approach (hybrid approach) was introduced by Case and Quigley (1991) and combines the desirable aspects of the hedonic and repeat-sales measures. The combination of the hedonic and repeat-sales information updates the hedonic regression's error structure to improve the efficiency of estimation of the price index and hedonic parameters (Quigley, 1995).

The building block for the hybrid model is the pooled hedonic regression (1) restated for convenience in vector notation:

$$p_h^t = \alpha + \boldsymbol{\beta}'\mathbf{x}_h^t + \boldsymbol{\gamma}'\mathbf{d}_h^t + \varepsilon_h^t.$$

$\boldsymbol{\beta} = \{\beta_k\}$ is a $K \times 1$ vector of implicit prices and $\mathbf{x}_h^t = \left\{x_{k,h}^t\right\}$ a $K \times 1$ vector of characteristics observations for each house h. $\boldsymbol{\gamma}$ is the $T \times 1$ vector of price changes and $\mathbf{d}_h^t$ the $T \times 1$ vector of time-dependent dummy variables for each house h.

The key to understanding the contribution of the hybrid approach is to consider repeat-sales observations as non-independent and to separate out the hedonic regressions for single-sales $i = 1, \ldots, I$ and repeat-sales $j = 1, \ldots, J$:

$$p_i^t = \alpha + \boldsymbol{\beta}'\mathbf{x}_i^t + \boldsymbol{\gamma}'\mathbf{d}_i^t + \eta_i + e_i^t \tag{3a}$$
$$p_j^t = \alpha + \boldsymbol{\beta}'\mathbf{x}_j^t + \boldsymbol{\gamma}'\mathbf{d}_j^t + \eta_j + e_j^t \tag{3b}$$
$$p_j^\tau = \alpha + \boldsymbol{\beta}'\mathbf{x}_j^\tau + \boldsymbol{\gamma}'\mathbf{d}_j^\tau + \eta_j + e_j^\tau. \tag{3c}$$

Note here the error term for house $h = \{i, j\}$ is decomposed into a time-independent specification error η_h and a white noise process e_h^t; in the above hedonic regression the error was expressed as ε_h^t. The hybrid contribution comes in differencing equation (3b) from equation (3c) to yield the following system of equations:

$$p_i^t = \alpha + \boldsymbol{\beta}'\mathbf{x}_i^t + \boldsymbol{\gamma}'\mathbf{d}_i^t + \eta_i + e_i^t \tag{4a}$$
$$p_j^t = \alpha + \boldsymbol{\beta}'\mathbf{x}_j^t + \boldsymbol{\gamma}'\mathbf{d}_j^t + \eta_j + e_j^t \tag{4b}$$
$$p_j^\tau - p_j^t = \boldsymbol{\gamma}'\boldsymbol{\delta}_j^\tau + e_j^\tau - e_j^t, \tag{4c}$$

where $\boldsymbol{\delta}_j^\tau$ is a $T \times 1$ vector of time-dependent variables equal to 1 for the second of the pair of sales for house j, -1 for the first sale of the pair and 0 otherwise. Here, the error structure is updated as the unobserved specification errors are removed using repeat-sales observations. Assuming, for house $h = \{i, j\}$, $\mathsf{E}(e_h^t) = 0$, $\mathsf{Var}(e_h^t) = \sigma_e^2$ and $\mathsf{Cov}(e_i^t, e_j^t) = \mathsf{Cov}(e_h^t, e_h^\tau) = 0$, $\mathsf{E}(\eta_h) = 0$, $\mathsf{Var}(\eta_h) = \sigma_\eta^2$ and $\mathsf{Cov}(e_h^t, \eta_j) = 0$, the $(I + 2J) \times (I + 2J)$ covariance matrix of the error structure is:

$$\Omega = \begin{bmatrix} (\sigma_\eta^2 + \sigma_e^2)\mathbf{I} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & (\sigma_\eta^2 + \sigma_e^2)\mathbf{I} & -\sigma_e^2\mathbf{I} \\ \mathbf{0} & -\sigma_e^2\mathbf{I} & 2\sigma_e^2\mathbf{I} \end{bmatrix}. \tag{5}$$

Note that with the error assumptions we can write $\sigma_\varepsilon^2 = \sigma_\eta^2 + \sigma_e^2$ where σ_ε^2 is the variance of the error of regression (1). Conducting feasible generalised least squares on equation system (4a), (4b) and (4c) with an estimate of the covariance matrix (5) constitutes the hybrid hedonic repeat-sales approach used in this paper.

4.1 Estimation Procedure

An estimation procedure for the hybrid model is outlined below. We assume misspecification errors are uncorrelated with the included variables which opens up an interesting comparison between the hybrid model and the random effects model. Consistent estimators of components of the covariance matrix are based on the analysis of the random effects model by Greene (2003).

To derive $\hat{\sigma}_e^2$ a repeat-sales regression (2) for the time period analysed is conducted. Let the $n = 1, \ldots, N_{\mathrm{RS}}$ residuals of this model be $\hat{\xi}_n$. From the error term of (4a), $\sigma_e^2 = \frac{1}{2}\mathsf{Var}(\xi_n)$. Using the estimated residuals, an unbiased and consistent estimator with the degrees of freedom adjustment is (applying Greene, 2003, p. 297):

$$\hat{\sigma}_e^2 = \frac{1}{2}\left(\frac{1}{N_{\mathrm{RS}} - T}\right)\sum_{n=1}^{N_{\mathrm{RS}}} \hat{\xi}_n^2. \tag{6}$$

It can be proved that the estimate of $\hat{\sigma}_e^2$ is consistent (not done here). The estimation of $\sigma_\varepsilon^2 = \sigma_\eta^2 + \sigma_e^2$ is more straightforward. Here, applying Theorem 10.8 in Greene (2003, p. 210) stating that the ordinary least squares estimators are consistent asymptotically, the ordinary least squares residuals $\hat{\varepsilon}_n$ of a hedonic regression (1) on all $n = 1, \ldots, N$ observations, single and repeated, may be used:

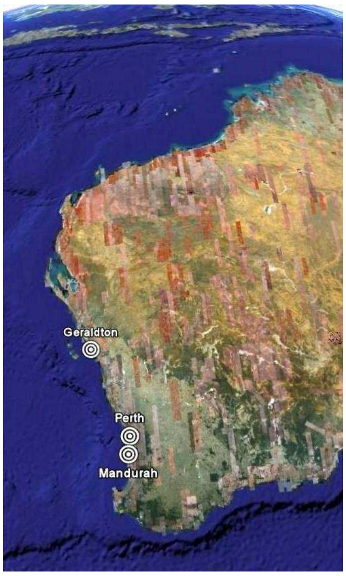
$$\hat{\sigma}_\varepsilon^2 = \left(\frac{1}{N-K-T-1}\right)\sum_{n=1}^{N}\hat{\varepsilon}_n^2. \qquad (7)$$

Using $\hat{\sigma}_e^2$ and $\hat{\sigma}_\varepsilon^2$, an estimate of covariance matrix (5) is formed, and feasible generalised least squares estimators of the hedonic and time dummy variables are estimated.

5 Western Australian House Prices

This section investigates the relative performance of the hybrid approach outlined above. Land titles office data on house sales in Geraldton, Mandurah and Perth were generously provided by the Real Estate Institute of Western Australia (REIWA) for the analysis. Figure 1 displays the location of Geraldton, Mandurah and Perth in Western Australia.

Figure 1: Location of Geraldton, Mandurah and Perth



Source: Google Earth (2008)

The data were filtered for anticipated measurement and recording errors. Duplicate observations were removed, as were houses not classified as individual sales. Properties that sold for less that \$50,000 were filtered out to control for transactions not conducted between independent parties. Further, houses selling for \$10,000,000 or more were excluded to control for data entry errors. For example, one filtered Geraldton house was recorded with an unrealistic, but curiously accurate price of $30,922,010. To maintain consistency with REIWA median price calculation methodology, houses greater than one hectare in land size were removed, as were those less than 200m^2. Reid (2008) made a similar adjustment.

The dataset was detailed in its characteristics observations. From the raw variable characteristics, a number of hedonic variables were formed, by considering the 'service' provided by each recorded house characteristic and with the goal of parsimoniously representing the hedonic regression equation. The raw variables and constructed hedonic variables are listed in Table 2. A few notes need to be made. First, 'Bedrooms' consists of bedrooms and studies, as these are relatively substitutable rooms (Reid, 2008). Second, 'Recreation Rooms' combines lounge, family and games rooms and third, 'Culinary Rooms' comprises of meals and dining rooms. Carports and garages, which were recarded as attached (ATT), detached (DET) and 'under the main roof' (UMR) were aggregated into 'Carparks'.

Table 2: Raw and Hedonic Variables

Raw Variables		
Sale Price	Bathrooms	Pool (Type)
Sale Date	Ensuites	Tennis Courts
Suburb	Kitchens	Carports (ATT)
Land Area	Dining Rooms	Carports (DET)
House Area	Meals Rooms	Carports (UMR)
Year Built	Games Rooms	Garage (ATT)
Wall Type	Lounge Rooms	Garage (DET)
Roof Type	Family Rooms	Garage (UMR)
Bedrooms	Study Rooms	Zoning
Hedonic Variables		
Sale Price	Bedrooms	Carparks
Sale Date	Bathrooms	Pools
Wall Type	Kitchens	Tennis Courts
Roof Type	Culinary Rooms	Land Area
Year Built	Recreation Rooms	Distance from CBD

A geographic variable which is thought to influence the price of a house is the distance of the house from a city's central business district (CBD). In this analysis, this was considered to be a determinant of Perth house prices as opposed to Geraldton and Mandurah prices based on their relative sizes. To determine the distance of a house from the Perth CBD a unique location for each house was required. Unfortunately, the dataset did not provide these explicitly; as noted in Table 2, the finest locational record in the datasets was at the suburb level.

To approximate each house's unique location, suburb coordinates were drawn from the spatial software program GOOGLE EARTH (2008) and each house in that suburb was assigned the suburb coordinate plus a normally distributed random component. The random component was designed to ensure 99% of the observations fell within 1km of the suburb coordinate. The idea of the procedure was to ensure houses close to each other record a unique and reasonably accurate location. It should also be noted individual house coordinates allowed the testing of, and correcting for spatial autocorrelation where required.

After constructing each unique location, distances were calculated between the Perth CBD and each house by applying spherical trigonometry which accounts for the gentle curvature of the globe (Geoscience Australia, 2008). Each of the sub-markets were divided into the groups North, South, East and West from the CBD and the resulting four variables included in quadratic form in the Perth hedonic regression.

5.1 Models Constructed

This section details the models constructed for each dataset which requires a consideration of the size of the dataset and the relative merits of each model.

Geraldton

The filtered Geraldton dataset consisted of 7,796 observations. Considering the relatively small size of the dataset, all observations are combined over all time periods 1994 to 2007 and a pooled hedonic model, a pooled repeat-sales model and a pooled hybrid hedonic repeat-sales model were constructed. The median and arithmetic mean were formed for comparison to these regression models. A mix-adjusted measure was not constructed in consideration of the sample size, though conceivably one could be formed by stratifying suburbs into SEIFA Index quartiles.

Mandurah

After filtering, 22,958 observations spanning the fourteen years 1994 to 2007 remained for the Mandurah analysis. In a similar issue for the Geraldton dataset, in consideration of the relatively small size of the dataset, pooled hedonic, repeat-sales and hybrid hedonic repeat-sales models are constructed; a model based on the mix-adjusted methodology was omitted. The arithmetic mean and simple median were formed for comparison.

Perth

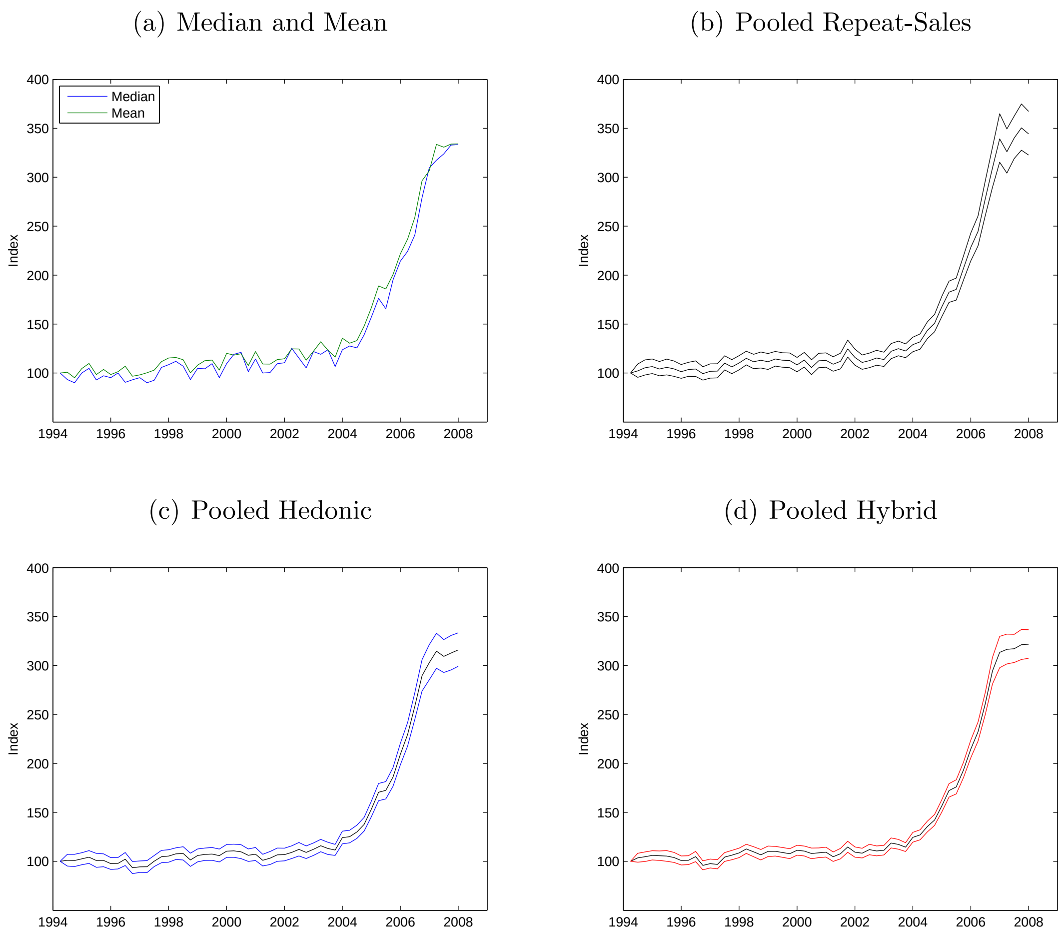
The size of the Perth dataset (404,767 filtered observations) allows more flexibility in the creation of house price indexes. An unweighted mix-adjusted measure based on 22 REIWA-defined “sub-regional areas” was formed. Pooled hedonic, repeat-sales and hybrid hedonic repeat-sales models are also constructed and presented. Further, to respond to potential criticism that pooling house sales observations across 14 years is too restrictive on the parameter coefficients, an adjacent period hedonic price index is formed by conducting hedonic regressions across successive quarters. This procedure permits maximum parameter flexibility across time for the quarterly time-dummy price index.

5.2 Estimated Price Indexes

Geraldton

Figure 2 displays the five constructed price indexes for Geraldton with confidence intervals for the regression measures. All five show show similar overall price movement–relatively flat house prices followed by strong price appreciation from approximately 2004 with growth of around 230% from 1994 to 2007. In response, there has been a 'stream of proposed [development] projects' in Geraldton, outlined in detail by Saunders (2008*a*; 2008*b*).

Figure 2: Geraldton Price Indexes, 1994-2007



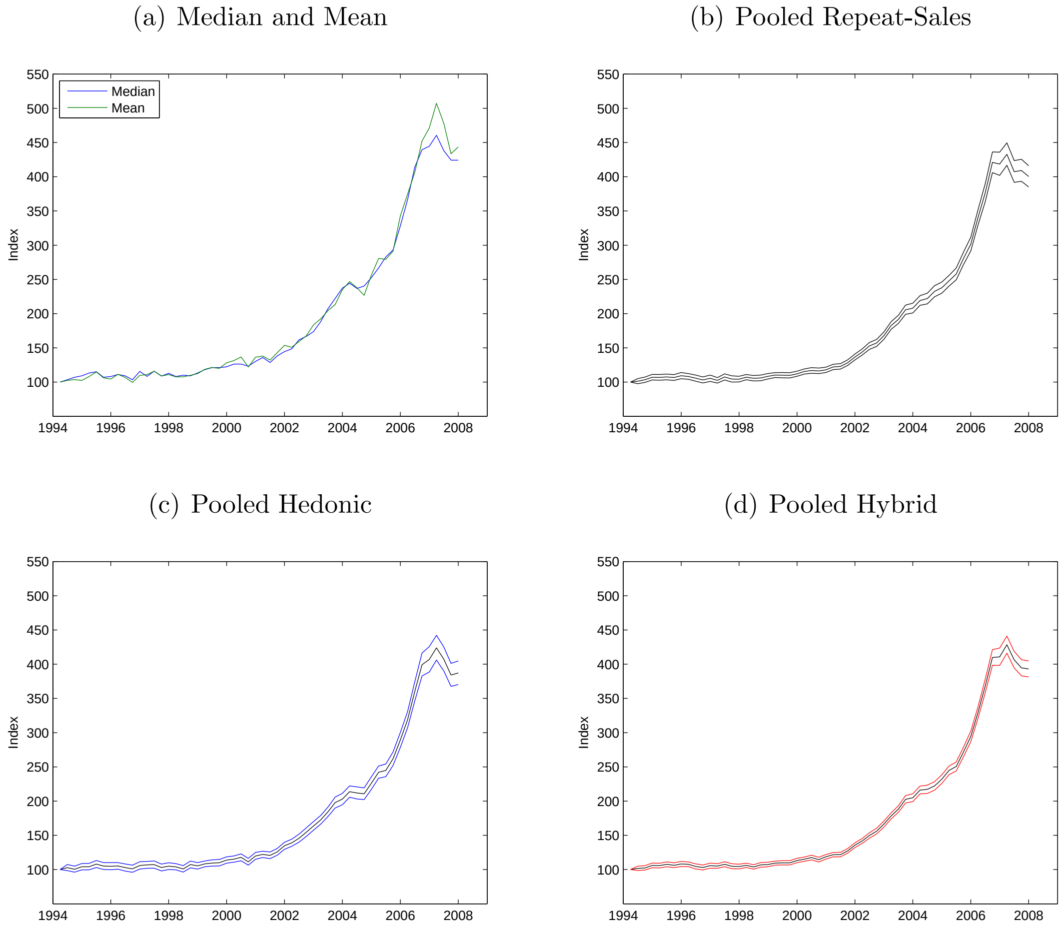
Note: The top and bottom solid lines in Panels (b) (c) and (d) give the 95% confidence bound about the constructed price index (the middle, black solid line).

A more detailed evaluation of the indexes (including the Mandurah and Perth price indexes) follows below, however 'eyeballing' Figure 2 suggests the regression methods are less volatile than the median measure (as expected) and that the hybrid model displays the narrowest confidence band about the constructed price index. This result is in line with the hybrid model exposition above; it augments the single-sale information with the data contained within repeat-sales pairs, resulting in more efficient price index estimators.

Mandurah

The five constructed indexes are presented in Figure 3. All show substantial price appreciation over the time period, ranging from an estimate of 284% increase (by the hybrid measure) to a 343% increase (according to the arithmetic mean). The strong growth in Mandurah house prices perhaps reflects recent significant population growth. However, in an indication that prices have risen too far, substantial declines in Mandurah house prices have been recorded throughout 2007–evidence against Case's (1986) remark that nominal house prices are sticky downwards.

Figure 3: Mandurah Price Indexes, 1994-2007



Note: The top and bottom solid lines in Panels (b) (c) and (d) give the 95% confidence bound about the constructed price index (the middle, black solid line).

Perth

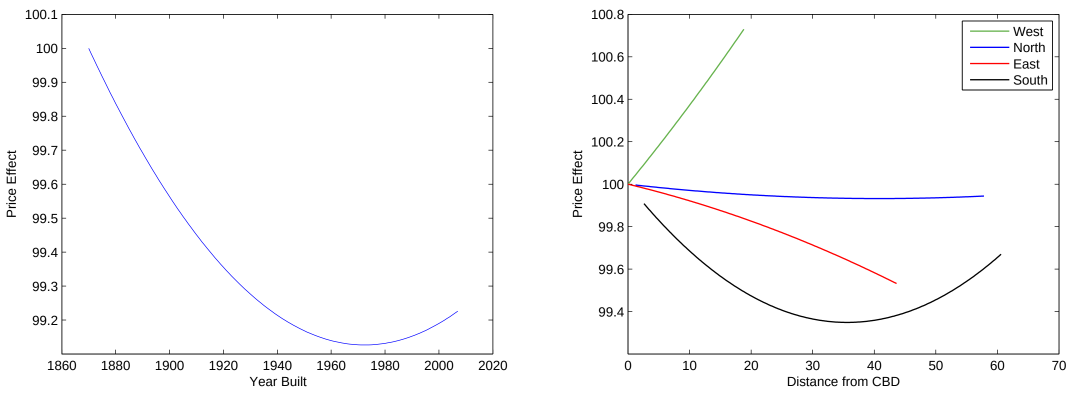
The regression output for the Perth hedonic regression is given in Table 3. An analysis of the hedonic variable coefficients is useful. First, the land area coefficient suggests that the amount of land has, as expected, a positive influence on the price of a house in Perth. Similarly, having a pool or a tennis court has a positive effect on house prices. Most other variables display the expected sign and gradient.

In addition, some implied figures can be derived from the regression coefficients. Figure 4(a) gives the implied quadratic shape of the effect of the 'year built' variable on price for the minimum to the maximum observed year built in the Perth dataset. The interesting and intuitive result is the turning point of the quadratic function at a year built of 1972. Below this figure, houses are observed as having increasing 'heritage' value while above, increasing 'freshness' value.

Figure 4(b) shows the implied quadratic effect of distance from the Perth CBD on price for houses in the four rough cardinal quadrants. Noticeably, heading in different directions from the CBD has widely varying effects on price; in the direction West, as houses move closer to the beach there is a strong positive effect on price. If we move North along the Indian Ocean coastline, there is little effect of distance on price, while if we head East there is an implied negative effect on house prices which does not turn up. Moving South, there is at first an implied decline in house prices due to distance, before becoming positive around 35km from the CBD.

Figure 4: Perth, Effect of Variables on Price

(a) Year Built on Price (b) Distance on Price



Note: This figure shows the constructed effect of a house's year built and distance characteristics on price, derived by forming the price effect from the quadratic hedonic coefficients of the OLS regression; the x-values for the quadratic function are bounded by the minimum and maximum year built and distance values in the Perth dataset.

The seven constructed price indexes are presented in Figure 5. Narrow confidence intervals, though difficult to identify are displayed on the indexes of the regression methods. Again, we see very strong price growth after 2000; over the entire period 1994 to 2007, Perth house prices grew by 283% (by the hybrid method).

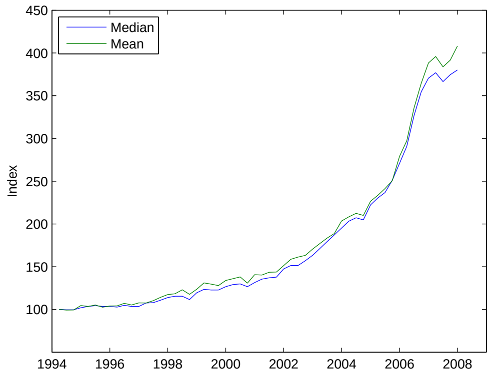
Table 3: Perth Pooled Hedonic Regression, Ordinary Least Squares

$$\ln p_h^t = \alpha + \sum_k \beta_k x_{k,h}^t + \sum_\tau \gamma^\tau D_h^\tau + \varepsilon_h^t$$

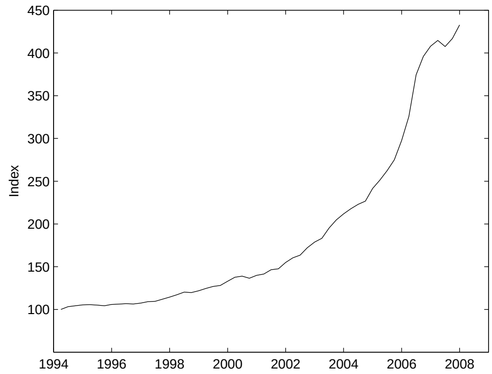
Variable	Coefficient	t-statistic	p-value
Constant	307.282	69.074	0.0000
ln Land Area	0.273	199.809	0.0000
Pool	0.082	55.815	0.0000
Tennis Court	0.427	18.298	0.0000
Year Built	−0.301	−66.349	0.0000
Year Built^2	0.000	66.074	0.0000
Distance North	−0.003	−6.479	0.0000
Distance North^2	0.000	5.228	0.0000
Distance East	−0.007	−10.300	0.0000
Distance East^2	−0.000	−4.551	0.0000
Distance South	−0.037	−47.295	0.0000
Distance South^2	0.001	54.292	0.0000
Distance West	0.035	9.237	0.0000
Distance West^2	0.000	1.186	0.2351
Wall Type			
- Asbestos	−0.097	−40.156	0.0000
- Brick-Venere	−0.175	−61.934	0.0000
- Brick (base)	–	–	–
- Timber Frame	−0.034	−4.015	0.0001
- Weatherboard	−0.146	−39.404	0.0000
- Stone	0.144	11.586	0.0000
Roof Type			
- Asbestos	0.039	14.251	0.0000
- Iron	0.066	27.755	0.0000
- Tile (base)	–	–	–
- Aluminium	0.025	1.210	0.2262
- Metal	0.023	1.255	0.2093
Bedrooms			
- Bed1	0.007	0.885	0.3760
- Bed2	0.014	8.127	0.0000
- Bed3 (base)	–	–	–
- Bed4	0.046	34.773	0.0000
- Bed5	0.178	92.386	0.0000
- Bed6	0.246	55.445	0.0000
- Bed7	0.231	19.013	0.0000
Bathrooms			
- Bath1 (base)	–	–	–
- Bath2	0.144	97.629	0.0000
- Bath3	0.411	105.800	0.0000
- Bath4	0.642	62.267	0.0000
Culinary Rooms			
- Culinary0	−0.038	−25.921	0.0000
- Culinary1 (base)	–	–	–
- Culinary2	0.074	59.023	0.0000
- Culinary3	−0.008	−0.351	0.7257
Recreation Rooms			
- LFGRoom1 (base)	–	–	–
- LFGRoom2	0.018	13.929	0.0000
- LFGRoom3	0.014	7.606	0.0000
Carparks			
- CarPark0 (base)	–	–	–
- CarPark1	0.028	20.470	0.0000
- CarPark2	0.142	91.143	0.0000
- CarPark3	0.183	53.987	0.0000
- CarPark4	0.219	33.045	0.0000
Sub-Market Dummies			
Time Dummies			

Figure 5: Perth Price Indexes, 1994-2007

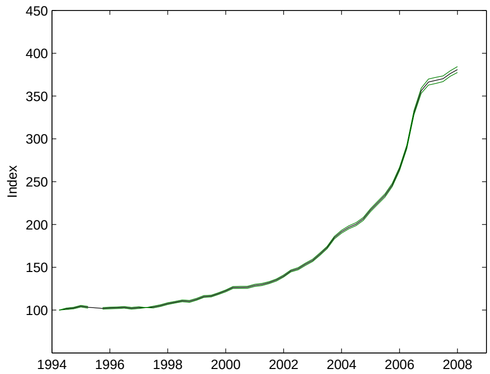
(a) Median and Mean



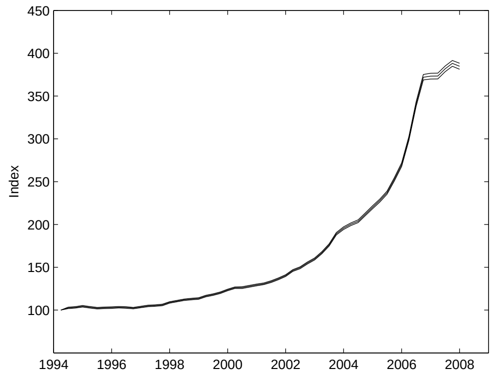
(b) Mix-Adjusted



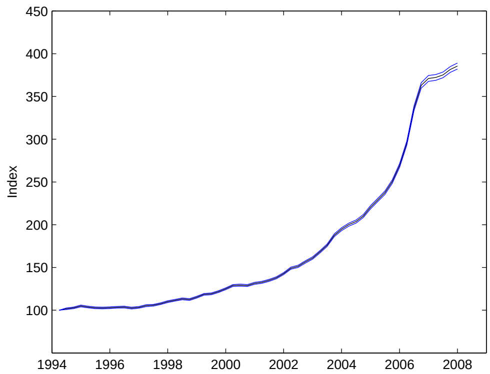
(c) Adjacent-Period Hedonic



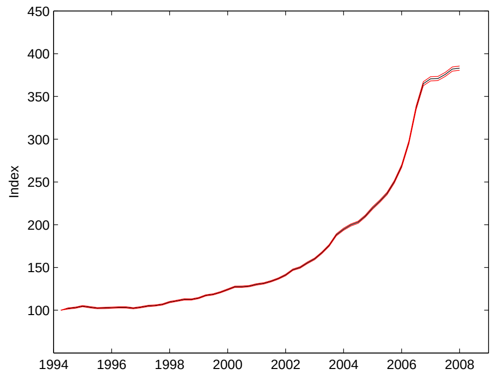
(d) Pooled Repeat-Sales



(e) Pooled Hedonic



(f) Pooled Hybrid



Note: The top and bottom solid lines in Panels (c), (d), (e) and (f) give the 95% confidence bound about the constructed price index (the middle, black solid line).

Again, it is not immediately clear from the graphs how the indexes perform relative to each other. However, it is noticeable the regression methods provide a smoother price progression relative to the median and mix-adjusted measures. It is also noticeable that the regression methods are very similar to one another, an observation which will be discussed in the section below, which analyses the indexes in detail and relates the results to the model estimation procedure detailed above.

5.3 Price Index Evaluation

Here, the different constructed price indexes are compared and analysed. As has already been observed, a scan of the price indexes and their confidence intervals suggests the hybrid model gives the most precise estimators of the regression methods. Building on this observation, Table 4 provides additional details of the calculated price indexes including statistics on the estimated overall growth rate, the average annual growth rate, the average quarterly growth rate, the quarterly growth rate standard deviation to gauge index volatility and the average confidence interval width for the regression methods.

The first point to note is the relative volatility of the indexes. From column (5), the regression methods display a smaller index volatility relative to the mean and median, particularly in the smaller datasets. Second, the regression methods, in most cases displayed a smaller overall growth rate compared to the mean and median, reflecting positive quality changes. However, for Perth, the median gave lower overall growth; this may reflect the composition of houses in the sample and is not inconsistent with observed quality improvements because the sample composition controlling mix-adjusted measure gives much higher overall growth compared to the regression methods.

Based on confidence interval widths, Table 4 suggests the repeat-sales method performs reasonably poorly with a small dataset confirming Meese and Wallace's (1997) observation that the repeat-sales methodology is sensitive to small samples. For the Geraldton dataset, the confidence interval width about the constructed repeat-sales index was wider than the other regression methods. For the larger datasets, the repeat-sales method performed better than the pooled hedonic index.

The time-dependent dummy variable has the advantage of providing confidence intervals about the price index. As for any price index involving uncertainty, this is a desirable analytical device (Clements, Izan and Selvanathan, 2006). Obviously, the tighter the confidence band about the estimated price index, the more precise the estimate. Column (6) of Table 4 gives the average confidence interval width for the available regression methods. As indicated in the table, for each of the datasets, the constructed pooled hybrid hedonic repeat-sales model produced price index estimates with the tightest confidence bands; we can be more confident in their precision. These empirical results accord with the theory outlined above.

However, it is relevant to note the small width of the Perth confidence bands. Despite the results suggesting the hybrid measure provides the most precise estimator,

Table 4: Price Index Statistics, 1994-2007

Measure (1)	Growth Statistics (%) Overall (2)	Annual (3)	Quarter (4)	Std Dev (5)	CI Width (6)
Geraldton					
- Mean	234.1	8.994	2.480	7.367	–
- Median	233.3	9.262	2.599	8.881	–
- Repeat-Sales	244.2	8.721	2.394	5.062	19.453
- Pooled Hedonic	215.9	8.191	2.219	4.707	15.869
- Pooled Hybrid	225.7	8.189	2.262	4.433	13.117
Mandurah					
- Mean	343.6	11.065	2.918	5.987	–
- Median	324.2	10.378	2.775	4.847	–
- Repeat-Sales	300.3	10.104	2.623	3.819	12.810
- Pooled Hedonic	287.8	9.978	2.581	4.254	14.819
- Pooled Hybrid	284.9	9.850	2.556	3.967	11.385
Perth					
- Mean	308.1	10.855	2.650	3.563	–
- Median	280.1	10.457	2.503	3.113	–
- Mix-Adjusted	332.8	11.132	2.737	2.824	–
- Adjacent Hedonic	288.1	10.498	2.530	2.718	3.036
- Repeat-Sales	284.8	10.523	2.519	2.871	2.797
- Pooled Hedonic	285.5	10.426	2.518	2.730	3.014
- Pooled Hybrid	283.1	10.423	2.508	2.760	2.170

given the small width, the gain in precision is relatively small. This was, at first, an interesting observation; it suggests that as the dataset increases in size, the regression methodology applied becomes increasingly irrelevant. Hansen (2006, p. 23) made a similar remark when reviewing hedonic and repeat-sales indexes for Sydney, Melbourne and Brisbane. However, for the smaller datasets of Mandurah and Geraldton, the hybrid method did produce noticeably more precise estimators.

The availability of confidence intervals also allows an observer to identify whether movements in the price index differ (statistically) significantly from the previous time period by observing whether the price index in the current time period lies outside the estimated confidence band for the previous time period. This provides a much more robust interpretation of the price index and its movements, particularly for housing which is subject to compositional and quality biases.

Table 5 details the results obtained by taking this approach to the analysis of the indexes. It gives answers to two questions regarding the regression method indexes: (i) does the median index lie within the regression method confidence interval? (Panel *A*); and (ii) does the median measure provide a statistically different growth rate from the regression methods? (Panel *B*). The second question is answered by reference to what is defined here as the 'plausible growth rate' for the regression indexes; the upper bound of the plausible growth rate range is the difference between the upper estimated confidence interval in the future period and the lower estimated

Table 5: Price Indexes: Median and Regression Measures

	Geraldton		Mandurah		Perth	
Measure	Number	%	Number	%	Number	%
A. Median Within Confidence Bounds?						
Adjacent Hedonic	–	–	–	–	55	100.00
- Yes	–	–	–	–	19	34.55
- No	–	–	–	–	36	65.45
Pooled Hedonic	55	100.00	55	100.00	55	100.00
- Yes	40	72.73	7	12.73	23	41.82
- No	15	27.27	48	87.27	32	58.18
Pooled Hybrid	55	100.00	55	100.00	55	100.00
- Yes	29	52.73	8	14.55	9	16.36
- No	26	47.27	47	85.45	46	83.64
Repeat-Sales	55	100.00	55	100.00	55	100.00
- Yes	28	50.91	13	23.64	9	16.36
- No	27	49.09	42	76.36	46	83.64
B. Median Growth Within Plausible Range?						
Adjacent Hedonic	–	–	–	–	55	100.00
- Yes	–	–	–	–	39	70.91
- No	–	–	–	–	16	29.09
Pooled Hedonic	55	100.00	55	100.00	55	100.00
- Yes	49	89.09	55	100.00	41	74.55
- No	6	10.91	0	0.00	14	25.45
Pooled Hybrid	55	100.00	55	100.00	55	100.00
- Yes	38	69.09	52	94.55	29	52.73
- No	17	30.91	3	5.45	26	47.27
Repeat-Sales	55	100.00	55	100.00	55	100.00
- Yes	46	83.64	54	98.18	31	56.36
- No	9	16.36	1	1.82	24	43.64

confidence interval in the current period, while the lower bound of the plausible growth rate range is the difference between the lower estimated confidence bound in the next period and the upper estimated confidence bound in the current period. This question was included because it was thought answers to the first question depend on the prior realisations of each price index; that is, the growth rate is specific to two time periods only.

The following conclusions may be drawn from Table 5. First, for Geraldton and Perth the hybrid model provides a higher proportion of statistically different price indexes estimators when compared to the median relative to the corresponding comparison between the hedonic and median price indexes. Indeed, for Geraldton, 73% of the constructed median price index measures fall within the confidence bounds of the pooled hedonic price index which suggests there is little advantage to using the hedonic approach apart from an apparent smoothing of the price index.

Second, from Panel B the constructed median growth value is more likely to lie outside the possible plausible range of the hybrid measure relative to the pooled hedonic measure. This is likely to reflect the greater precision possible in estimating these price indexes using the hybrid approach. The apparent high proportion of the median growth rate falling within the plausible range should be tempered by the fact the plausible range is constructed using the two upper-bounds of successive constructed price index confidence intervals; the possibility of this realisation is therefore small. In sum, these conclusions support the precision benefit of the hybrid model relative to the hedonic method.

6 Conclusion

This paper has summarised the methodology and results of an analysis of methods of modelling house price movements. It began with with some motivating comments on why precise estimates of house price changes are desirable. In short, consumers and potential home owners need accurate measures to appropriately assess wealth levels and analyse the potential risks in purchasing a house. The paper examined the pros and cons of various measures of house prices, including the median and mix-adjusted measures, and the repeat-sales and hedonic regression methods. It provided an exposition of the analytically superior hybrid hedonic repeat-sales measure. Empirical results using Geraldton, Mandurah and Perth house sales data showed the greater precision of the hybrid measure relative to the hedonic and repeat-sales approaches, and illustrated the strong house price growth in Western Australia from 2002.

References

Bhatia, K. (1987), 'Real Estate Assets and Consumer Spending', *Quarterly Journal of Economics* **102**, 437–444.

Campbell, J. Y. and Cocco, J. F. (2007), 'How Do House Prices Affect Consumption? Evidence from Micro Data', *Journal of Monetary Economics* **54**, 591–621.

Case, K. E. (1986), 'The Market for Single-Family Homes in the Boston Area', *New England Economic Review* **May**, 38–48.

Case, K. E. and Quigley, J. M. (1991), 'The Dynamics of Real Estate Prices', *Review of Economics and Statistics* **73**, 50–58.

Case, K. E., Quigley, J. M. and Shiller, R. J. (2001), Comparing Wealth Effects: The Stock Market Versus the Housing Market, Discussion Paper 1335, Cowles Foundation for Research in Economics, Yale University.

Clements, K. W., Izan, I. H. Y. and Selvanathan, E. A. (2006), 'Stochastic Index Numbers: A Review', *International Statistical Review* **74**(2), 235–270.

de Haan, J. (2007), Hedonic Price Indexes: A Comparison of Imputation, Time Dummy and Other Approaches, Technical Report, Statistics Netherlands.

Dvornak, N. and Kohler, M. (2003), Housing Wealth, Stock Market Wealth and Consumption: A Panel Analysis for Australia, Research Discussion Paper 7, The Reserve Bank of Australia.

Engelhardt, G. (1994), 'House Prices and the Decision to Save for Down Payments', *Journal of Urban Economics* **36**, 209–237.

Engelhardt, G. (1996), 'House Prices and Home Owner Saving Behavior', *Regional Science and Urban Economics* **26**, 313–336.

Geoscience Australia (2008), 'Calculating Distance Between Two Points', http://www.ga.gov.au/geodesy/datums/distance.jsp/.

Google (2008), 'Google Earth', http://www.earth.google.com/.

Greene, W. H. (2003), *Econometric Analysis*, international fifth edn, Prentice Hall.

Hansen, J. (2006), Australian House Prices: A Comparison of Hedonic and Repeat-Sales Measures, Research Discussion Paper 2006-03, Reserve Bank of Australia.

Hansen, J. (forthcoming, 2009), 'Australian House Prices: A Comparison of Hedonic and Repeat-Sales Measures', *Economic Record* .

MacFarlane, I. (2004), 'House of Representatives Standing Committee on Economics, Finance and Public Administration'. Hansard.

Matusik, M. (2008), Median Prices, The Matusik Insight 362, Matusik Property Insights.

Meese, R. and Wallace, N. (1997), 'The Construction of Residential Housing Price Indices: A Comparison of Repeat-Sales, Hedonic-Regression and Hybrid Approaches', *Journal of Real Estate Finance and Economics* **14**, 51–73.

Prasad, N. and Richards, A. (2006), Measuring Housing Price Growth: Using Stratification to Improve Median-based Measures, Research Discussion Paper 2006-04, Reserve Bank of Australia.

Quigley, J. (1995), 'A Simple Hybrid Model for Estimating Real Estate Price Indexes', *Journal of Housing Economics* **4**, 1–12.

Reid, B. (2008), Hedonic Imputation House Price Indices: Bias and Other Issues, Working Paper 2008/1, School of Economics, The University of New South Wales.

Reserve Bank of Australia (2008), Selected Assets and Liabilities of the Private Non-Financial Sectors, Statistics Tables B20, Reserve Bank of Australia.

Saunders, C. (2008*a*), 'Boom Time as Northern Port City Bursts into Life', The West Australian.

Saunders, C. (2008*b*), 'Mid-West City Eyes Top Billing', http://www.thewest.com.au/.

Sheiner, L. (1995), 'Housing Prices and the Savings of Renters', *Journal of Urban Economics* **38**, 94–125.

Sinai, T. and Souleles, N. (2005), 'Owner-Occupied Housing as a Hedge Against Rent Risk', *Quarterly Journal of Economics* **120**, 763–789.

The Economist (2005), 'After the Fall'.