CONSTANT QUALITY HOUSE PRICES IN AN AUSTRALIAN CONTEXT

A CASE STUDY OF PORT PIRIE, SOUTH AUSTRALIA

Peter A. Rossini, B.App.Sc.(Val), Grad.Dip.Econ.Stats.(NE), Lecturer
Robin R. Kooymans, B.App.Sc.(PRM), Grad.Dip.Ec.(NE), AVLE(Ec), AAII, Lecturer
Paul J. Kershaw, M.Sc., Principal Lecturer

School of Economics, Finance and Property, Faculty of Business and Management
University of South Australia

Abstract
This paper examines the relationship between median and mean house prices, which are commonly misconstrued in the popular press to mean the asset price, and constant-quality house prices as measured by hedonic equations and repeat sales. Median and mean house prices for relatively homogeneous sub markets are also examined, to determine whether there is potential to gauge asset price movements without the time and effort involved in constructing repeat sale and hedonic price indices. The case chosen for examination is Port Pirie, in South Australia. This small city has had large apparent price increases over a period of time in which real prices of housing in many Australian regional areas have apparently fallen. The sales data is divided into more homogeneous submarkets, to ascertain whether mean or median closely reflect the asset prices in these sub markets. Hedonic price and repeat sales indices are then constructed so that the usefulness of mean and median prices as housing price movement indicators in Port Pirie can be ascertained.

Introduction
Newspaper articles and real estate reports frequently refer to measures of growth in the residential real estate market. This may vary from simple quotes suggesting huge increases in values based on a monthly increase in the median prices, to quotes of stagnant markets based on properties which have resold over a short period, to annual growth estimates based on time series analysis of mean or median prices. Some groups use econometric modelling or stratified methods to estimate growth based on “typical” properties.

The purpose of this paper is to consider some of the literature on this topic (mainly from the USA) and to use empirical testing to consider the validity of some of the methodologies. This paper is the starting point of a broad empirical study which will consider several hundred locations, but for this paper only one location will be considered. To this end the paper is likely to ask more questions than it answers.
The Use of Housing Price Indices in the USA

Housing price indices have been constructed in the United States of America for many years, most often at county, metropolitan or some lower level, on the implicit or explicit assumption that estimated price indices are only valid for quite small areas (Case et al. 1991). They have been considered essential to primary and secondary mortgage markets, producers of housing services, and consumers of owner occupied housing (Meese & Wallace, 1991).

“The most widely reported measure is the National Association of REALTORS quarterly publication of median sales prices of existing single-family homes from 119 metropolitan areas” (Meese & Wallace, 1991). The U.S. Bureau of the Census publishes an index of speculatively built homes, with figures taken on a sample basis (Nourse, 1990). The U.S. Bureau of the Census also conducts a Standard Metropolitan Statistical Area Annual Housing Survey of housing stock in sixty metropolitan areas in a three to four year cycle, which information is used to conduct tenure specific hedonic housing price indices based on location and dwelling quality (Thibodeau 1989).

Probably the most controversial example of the use of housing price indices in the United States of America is the Mankiw and Weil (1989) paper “the Baby Boom, The Baby Bust, and The Housing Market”, in which they suggested that the real price of owner occupied housing would decline by 47% between 1987 and 2007 because the Baby Bust generation was entering into its house buying years, supplanting the Baby Boom generation which had driven the market. This finding was taken up by the press and, among other effects, caused the share prices of the major secondary mortgage dealers to fall 20% in one day (Green, 1990). A large number of real estate academics criticised the Mankiw and Weil paper on a number of grounds, for instance: not allowing for the effect of housing quality and demographic influences other than age (Green, 1990); mis-specification of the model (Hendershott 1990, Hamilton 1990); methodology used (Holland, 1990); economists’ prejudices (Woodward, 1991); evidence from similar markets (Engelhardt and Poterba, 1990)..

There has been a strong industry lobby trying to tie FHA maximum loan ceilings to median house prices, to the extent that legislation was introduced to (but defeated in) the House of Representatives in 1989 (Hendershott & Thibodeau, 1990).

The point to be made here is that the construction of housing price indices and the implications of the trends that they disclose seem to matter very much in the United States, while they have received very little attention in Australia. Reading of the American literature reveals a great deal of academic attention to construction and interpretation of residential real estate price indices, whereas it does not seem to have received very much attention in Australia.

The Use of Price Indices in Australia

Housing statistics are collected, but with the notable exception of the Australian Bureau of Statistics, not much attention appears to be paid to the validity of the figures, their method of collection, or presentation. Perhaps this is due to the historical absence of a secondary mortgage market, the relative lack of sophistication of Australian lenders, and the tendency of many Australian property market participants not to conduct rigorous market research before participating.
The typical index is produced from sales data and uses median prices to indicate sale price “trends” with little apparent attempt being made to correct for quality. Such figures are published in the Press for the benefit of the public (Humphreys 1994, Seidel 1994), with caveats as to reliability which are often ignored in the editorial conclusions drawn. The real estate agency industry also publishes such figures in its journals (REI Journal). As few real estate industry participants are trained in the interpretation of statistics, it is probable that such figures are taken to be more indicative than they perhaps really are.

The Australian Bureau of Statistics publishes cost price indices for new (project) homes and, more relevant to this study, for established houses. In the latter case, there is explicit recognition of the variation in the movement of an index depending on physical characteristics and location and the problems of compositional change. The ABS stratifies the raw sales price data by geographic area and physical characteristics, then weights together the price movements in the individual strata. The procedure is carried out for a sample of Local Government Areas in each city. (ABS, 1994). While useful at a macro-level for economic policy-making, the restriction to individual capital cities and the absence of local level data means that this information is of limited use to market participants at the micro-level.

As Australia’s real estate market participants become better educated in the ways of the market, so it seems sensible to assume that demand for accurate, useful information will increase. At the very least, the real estate academic community probably owes it to the market to signal the dangers in taking the presently available information too literally.

**The Problems of Constructing Housing Price Indices**

The principal problem which arises is the availability of sufficient sales data, both in terms of amount and of comprehensiveness of information. If median or mean price is to be the only consideration, then price and time of sale are all that are needed. Even this can be a problem to obtain in some areas - not so much because the information is unobtainable, but because it is expensive to obtain. If the limitations of mean and median prices are taken seriously, then a far more detailed and comprehensive data base is required.

A related problem, once the data have been obtained, is the ability to extract and work with it in an economical and timely fashion. The advent of cheap, readily available and relatively powerful computing has helped a great deal here, but the prospective analyst still requires appropriate software and sufficient storage for what may be a very large quantity of data.

House price indices are normally constructed for a specific geographic area and a particular component of the housing stock (Case et al, 1991). This implicitly assumes that the index is valid for all the houses making up that component in that area. Sometimes this assumption is explicitly made in the publication of the data. To the extent that the validity of the assumption depends on careful selection of the area and the housing component, this will present the analyst with either a problem of model specification or a problem of bias in results.
An increase over time in the price of housing in a given area may reflect both the increase in the price of a given constant-quality house, and the increase in quality and quantity of housing bought in an area over that period. Thus it is possible for a price index to show (say) a 10% increase from period A to period B, while the price of a given constant-quality house actually declines. This might happen if houses sold in period A were mostly small and in poor condition, but the houses sold in period B were large and in good condition. The first requirement to overcome this is data relating to quantity and quality of housing, while the second is a suitable model to make use of the data, as discussed below.

**Mean and Median Prices**

The arithmetic Mean price of housing has been used for construction of indices but, because of the potential bias due to unusually low or high sale prices influencing the Mean, it has become more usual to quote the Median price. The findings of this paper show that there may not actually be a significant difference in Mean and Median prices if the data are able to be screened for aberrant entries. The use of the median is certainly preferable if the details of sales beyond price and time are unknown but, because of the quality differences mentioned above, it probably does not deserve the reliance which has tended to be placed on it by purveyors of price indices in Australia.

**Use of Relatively Homogeneous Submarkets to Adjust for Quality**

The principal issue here is the method of identification of relatively homogeneous submarkets. The literature suggests the use of hedonic equations, repeat sales, or a mixture of the two (e.g. Hendershott & Thibodeau 1991, Goodman 1978), as discussed below.

Another method, which was tried in this study and which is used by the S.A. Department of Environment and Natural Resources in rating valuation, is a priori identification of obvious submarkets. For instance, a typical outer-suburban project home subdivision, which caters for a first home buyer market is probably sufficiently homogeneous to minimise distortions in indices caused by different “ends” of the market selling well in successive time periods. The flaw in using this method alone is that the a priori assumption is untested, thus relying on the analyst having a very good insight into the market. This is not likely to be possible in construction of indices for many areas, or in the weighting of overall indices for the quality of housing sold in any period.

The use of statistical techniques such as Cluster Analysis is a further possibility for scientifically choosing sub markets. Such techniques do require a good deal of data, market experience and skill to be effective.

**Use of Hedonic Equations to Adjust for Quality**

The reliability of price indexes developed from hedonic regressions depends on correct model specification. As the hedonic equation results from interaction of demand for and supply of various housing characteristics, which will themselves vary across time and space, selection of functional forms must be approached from an empirical viewpoint (Palmquist, 1980). The empirical approach demands that sufficient reliable data on each sale be obtainable in useful form, to enable the experimentation necessary to construct the most appropriate hedonic model.
Derivation of the price index from the hedonic model in the semi-logarithmic form, containing data aggregated over more than one time period, may be carried out using dummy variables for period effects, with the coefficients of the dummy variables representing the average rate of change in price (Bailey et al, 1963). Palmquist (1980) questions the validity of aggregating data over more than one time period and suggests the usual alternative of modelling a separate hedonic regression for each period and calculating an index based on the estimated price changes of the various real estate characteristics for each period. Palmquist’s results were equivocal, and as he points out, the techniques should give identical results if the coefficients other than those attaching to time and the constant term remain the same between periods. In practice, and especially over many time periods, it is probable that the non-time coefficients will change, representing changes in consumer demand for various aspects of housing.

The United States studies appear to have the advantage of very large samples which, while available when constructing aggregate price indices over very populous areas, tend not to be available when examining price trends across the smaller markets existing in Australia. On balance, given the relatively small quarterly sample available in the study, the somewhat ambiguous nature of conclusions drawn from the literature and the relatively short time period covered by the study, it was decided to adopt the aggregated regression approach. The way is still open to experiment with the disaggregated approach in larger markets.

As well as the problem of model specification alluded to above, there is the obvious weakness that obtaining a complete set of reliable regressors will be difficult, if not impossible in practice (Case et al, 1991). The data sources which were used in this study are probably at the high end of the scale for comprehensiveness and reliability of publicly available data, especially when doubtful sales are removed. However, there is no doubt that more information on, say, interior condition, would enable specification of more robust and reliable hedonic models.

The hedonic equation may be further modified by using repeat sales as a sample (Case and Quigley, 1991) or using repeat sales as a “control” on regression results (Case et al, 1991). This is briefly discussed below.

**Use of Resales to Adjust for Quality**

The principle of using resales is theoretically quite simple - just take a property that has transacted at least twice during the study period, and work out the growth rate for the period between sales. Unfortunately, there are practical problems with this approach - problems which are unappreciated by some real estate agents who manage to derive a trend from one resale.

If it is possible to identify, in statistically significant numbers, resales involving arm’s length transactions of properties which have remained physically unchanged between sales, in markets whose preferences for the attributes provided by the resale properties have not changed, then the repeat sales method provides a very useful method of construction of price indices. In effect, resales of properties in which all variables except price and time of sale are constant, should provide a good basis for a reliable, explanatory simple regression model, or a distribution of sales whose trend can be analysed.
Of course, there is no such thing as a true repeat sale, because the property which has sold twice, with intervening period \( n \) months, is clearly \( n \) months older the second time around. Palmquist (1980) concluded that the depreciation aspect was important, and that repeat sales models should be adjusted for a depreciation coefficient obtained by an aggregate multiple regression using age of the structure as a variable. Other researchers, notably Case et al (1991), find that depreciation is not a significant issue in some markets. Given the relatively short period of the study in this case and the fact that age of the structure did not show up as a significant variable, depreciation was not taken into account.

Case et al (1991) tested a variety of hedonic, repeat sales and hybrid models on a large data set. They found that the results of repeat sales models varied greatly, depending on how much care was taken to remove those sales in which characteristics other than age and time of sale had varied. They found that a naive repeat sales model, in which all repeat sales were used regardless of provenance, to be the least reliable. Another interesting observation of this study was that aggregated repeat sales models showed smaller Standard Errors and narrower Confidence Intervals, as well as lower rates of price appreciation, than hedonic models. This could be because repeat sales models are estimated on the part of the market with the shortest time between transactions and thus the smallest variances in characteristics making up the error term, or it could be that resales form a distinct submarket. The regression models estimated solely on the resold properties also showed lower rates of price appreciation, tending to support the latter hypothesis.

**The Study Area - Port Pirie, South Australia**

The literature suggests that using simple methods to assess growth are most likely to be a problem in locations where housing has changed over time. This could occur for a wide variety of reasons including rapid gentrification resulting in upgrading of much of the housing, urban redevelopment programs or major subdivision or housing projects adding a new type of housing. For this study a location was chosen that had some elements of all three in a situation capable of isolation from surrounding locations. The study area is in Port Pirie some 230 kilometres north of Adelaide. The town has a population of about 15,000 (and falling) and is very isolated. Port Pirie is notable for having one of the world’s largest silver-lead-zinc smelters. This smelter has led to large sectors of the town suffering from lead contamination. In the mid 1980’s the South Australian government implemented a plan to work with the local council to rectify much of this decontamination and hopefully reduce the very high lead levels which several studies had discovered in Port Pirie residents.

At the time the decontamination program was introduced, much of the housing was cheap and very substandard. Prices varied from as little as $200 to as high as $140,000 in 1986. The first stage of the decontamination program involved buying of the cheaper (non-renovatable) houses, clearing the sites to a depth of at least 30 cm, replacing this with “clean” soil and selling the site as a vacant site. Then progressively over a 7 year period, other contaminated housing was assessed and de-contaminated where possible. Depending on the level of contamination, this ranged from the stripping of paint, floor coverings and removal of all dust, to major recladding and redevelopment of the properties.
The result of this was that the very bottom end of the market was completely removed and the next-to-bottom end improved. At the same time, in other parts of Port Pirie, several modern subdivisions were going ahead, adding large modern houses to the housing stock.

The Study Period
The study period is from January 1986 to December 1992. This period corresponds to the period when the majority of the decontamination was carried out. Where necessary, data was analysed on a quarterly basis. There are normally between 50 and 70 sales in each quarterly period.

Mean and Median House Prices in Port Pirie
The starting point of the analysis was to establish summary statistics for each quarter of the study period, using all sales in the town of Port Pirie. Sales were carefully selected, rejecting all sales where there was evidence of abnormality in the data which would suggest a non-market transaction. The resulting figures were very similar to (in many cases the same as) the published Mean and Median figures from the South Australian Department of Environment and Natural Resources (DENR). DENR is responsible for collecting the original data and uses a variety of tests to remove likely non-market transactions when compiling its summary data. Figure 1. show the time series graphs using the data for the whole of Port Pirie. The Mean and Median figures are very consistent, the Median always a little lower than the Mean as would be expected in a market dominated by low priced housing. The Standard Deviation is of the same magnitude indicating the wide range of prices. This is also reflected in the Maximum and Minimum values.

Figure 1.
Regression was used to estimate annual growth rates for each of the indicators. The high rate of increase in the Minimum highlights the removal of the bottom sector of the market during the study period. Noticeably, the growth rate using the Mean and Median are different by over 1 percent per annum. Thus while the graph shows that the trends are consistent, there is a marked difference in the indicated growth rate. The Trimmed Means show very much what would be expected. A small trim (5%) produces results very similar to the mean while a larger trim (10%) approaches the Median.

<table>
<thead>
<tr>
<th>Measure</th>
<th>Indicated Growth Rate</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mean</td>
<td>8.90%</td>
</tr>
<tr>
<td>Median</td>
<td>9.97%</td>
</tr>
<tr>
<td>Min</td>
<td>27.85%</td>
</tr>
<tr>
<td>Max</td>
<td>6.45%</td>
</tr>
<tr>
<td>5% Trim Mean</td>
<td>8.97%</td>
</tr>
<tr>
<td>10% Trim Mean</td>
<td>9.48%</td>
</tr>
</tbody>
</table>

A Stratified Approach to the Problem
Port Pirie is made up of five identifiable areas. One incorporates the business centre, smelter and major community facilities with very little housing. This area was ignored for this study as there are generally only one or two sales per quarter. The first two locations used are Solomontown, which was the first major housing area in Port Pirie, and Port Pirie West, located to the west and north of Port Pirie adjacent to the smelters. These two locations contain most of the poor housing which was badly lead contaminated. Risdon Park and Port Pirie South are located south and south-west of the smelter and business area and contain more modern housing and some new subdivisions. This area has a much lower percentage of contaminated houses. The data for the 28 quarters was stratified by these four clearly identifiable locations.

Figure 2. shows the chart for Solomontown. This chart shows some typical trends. The Mean and Median are much less consistent, while the Standard Deviation is relatively small compared to the Mean and Median. In each location the Mean was a more consistent figure over time.

![Solomontown - Price Trends](chart.png)
Trend lines for each location were developed using the Median prices. These are shown on figure 3. Each location has an easily identifiable price level. The most notable observation is the that the Total figure (for all sales in Port Pirie) fits neatly between all the locations and is clearly not “typical” of any part of Port Pirie.

The data was then stratified by housing type. Two types of housing were found to be prevalent in Port Pirie. Over 50% of the housing can be described as either 4 roomed cottages in poor to average condition or 5 or 6 roomed conventional houses in average to good condition. These two types of housing are therefore fairly representative of the typical or normal property in Port Pirie. Some literature suggests that the Median reflects the value of the typical house. Figure 4. shows the graphs of each of these types and the total for all houses. Contrary to this suggestion, the overall Median value is clearly not reflecting the result from either of the housing types most prevalent in Port Pirie.

Table 2. shows the indicated annual growth rates for the stratified samples using a log-price regression method.
The Hedonic Approach

For the hedonic approach two regression models were estimated. Each model used all of the sales which had occurred over the 7 year period. The log of price was used in order to estimate growth as a normal compounding annual growth rate. In the first model the independent variable was the period measured in quarters. In the second model, further quantitative variables were added to try to account for major locational variations in Port Pirie and major differences in housing types. In this model, the house area and condition were found to be major value determinants. Two of Port Pirie’s sectors, Solomontown and Port Pirie West were found to have a major negative influence in prices as were houses of Austerity or Conventional style or with iron clad walls. Solid wall construction was found to have a major positive influence on prices. The variable measuring time as quarterly periods was also a highly significant indicator.

The indicators from the two models provide different indications of growth. The simple model (not adjusted for quality variations) indicates an increase of 11.2% per annum. The multiple regression model including quantitative and qualitative adjusters indicates an increase of 6.98% per annum.

Using Re-Sales to Estimate Growth

Because of the relatively small amount of data, resales were analysed on an annual basis. Figure 5. shows the distribution of percentage annual growth rates using all resales found over the entire period. These resales are divided into bad sales - for which there is clear evidence that the property in question has changed substantially between the first and subsequent sales, e.g. those properties bought by the council, decontaminated, then sold - doubtful sales, for which there is a strong possibility that the property has been altered between sales - and good sales, for which there is no evidence that the property may have been significantly changed. Because the sales were not individually analysed, there is always the possibility that some of the “good” sales are not, however this is a practical data limitation with which any analyst of mass data must learn to live. The distribution shows a concentration of indicated annual growth rates of between 0 and 10%, but there is a clear warning in the distribution that naive use of a few resales to estimate market growth could produce spectacularly implausible results!

Figure 5.
The data were further refined by the removal of sales which took place within less than a year after the previous sale. This was felt necessary because the resale of property within a year is highly likely to be a speculative transaction or one in which the vendor is under pressure to sell due to an unforeseen event, and because a reasonable profit in dollar terms over, say, 3 months, will read as a very high annual percentage increase. The distribution of resales over one year is shown in Figure 6. Comparison of the distributions shows that the technique has removed substantial numbers of the bad and doubtful sales, with a slightly greater concentration of properties within the 0 to 10% band.

Table 3 provides a summary of the growth rates indicated by resale analysis. Using resales which are more than one year apart shows that the mean growth rate is apparently a good 2 percentage points above the median. Given that there are substantial numbers of sales with growth rates of 15% or more, while there are almost negligible numbers with growth rates of less than -15%, this is to be expected. There seems to be a case for the median figure here.
Figure 7. provides a cumulative distribution of indicated growth rates from resales, reflecting the leaning towards positive growth rates and the concentration between 0 and 10%. The marker indicate that 50% of the cases fall between an indicated growth rate of 0% and 10%.

Summary of the Methods

The commonly used method of indicating growth by using the median prices for all sales, after removing sales of obviously doubtful utility, produces an annual growth rate over the study period 1986-1992 of 9.97% in the median price. It was interesting to observe that the mean price growth, at 8.9% per annum, produced a slightly better indicator in view of the subsequent results of the hedonic and resales method.

Stratifying the sample into four suburbs shows a growth rate of 13.94% in Solomontown, one of the two suburbs most affected by the decontamination program. This appears to be a classic example of the problem of ascribing increased average prices to asset price growth, when what we saw in Solomontown and Port Pirie West (10.29%) was a very marked quality improvement leading to
higher prices. This is emphasised in Risdon Park, which was virtually unchanged in quality and which recorded growth of 3.71% per annum. A difference of more than 6% per annum between the published growth rate for Port Pirie and the actual growth rate for the suburb could cause potential investors to be misled. Stratifying the sample into the predominant housing types again demonstrates the fallacy of relying on aggregate price indices. The cottages, which formed the majority of houses which were decontaminated, appear to have obtained their growth through a quality improvement, whereas the prices of later-built conventional houses grew at an annual rate of less than half the index rate.

The hedonic approach, in which the model adjusted for quality differences, showed an overall investment growth rate of 6.98% per annum - nearly 3% per annum below the index rate. Within the limitations of the model, there is a clear indication of the need to adjust for quality - at least in the market studied.

Using the median price of “good” resales more than a year apart produced an annual growth rate of 7.22%, which is reasonably consistent with the results obtained using the hedonic model. The figures are too close to judge whether the resales do form a distinct submarket - this would best be carried out using a hybrid model of regression on resales which was not done in this study because it was intended that the usefulness of each distinct method be tested.

Conclusion

The research that has been carried out in the United States on the problems in the construction of price indices appears to have some validity in Australia. It is apparent that the problem of median or mean prices not allowing for quality change, or for different ends of the market selling at different times, does exist. The study of Port Pirie demonstrated that the problem can be quite significant. Further studies will be carried out on other parts of the South Australian housing market to determine the extent of the possible misleading tendencies of the housing price indices used.

Potential solutions to the problem, which were derived from the United States literature, were applied to the case study. They all showed considerable promise, in that they produced results which were consistent with our understanding of the study area. The challenge to be met in studies of other areas is to produce methods which can be consistently and reliably used with data which are not supplemented by personal acquaintance. To this end, the use of regression models, resales, or a hybrid thereof, shows much promise.
References