An analysis of the theory and application of house price indices.

Almost all analysis of the housing market is based on ability to measure and track house price movement by a one dimensional index. Housing provides a set of unique challenges when constructing such a measure. Among them are the idiosyncratic nature of property units, relatively low sales frequency, and frequently highly non-normal price distribution. Because of these challenges, a number of methods for constructing house price indices (HPIs) have developed. Each of these methods exhibits a unique set of properties that lead to varying degrees of accuracy when applied to real data.

Recently there have been some changes to the way that some institutions in New Zealand construct their price indices. It seems timely to conduct a comprehensive comparison of HPI methods that discuss their relative strengths and weaknesses in the context of the challenges associated with index construction in this market. This would involve conducting robustness tests for bias and distributional mismatch with reference to generated and real data sets to investigate the performance and limitations of each.

There are three widespread HPI construction methods that are commonly used in some form by industry and government around the world to track house price movements. They are the repeat sales estimator, hedonic price estimation and variations on a median price index. More recently, institutions in Denmark, Sweden, the Netherlands, and New Zealand have adopted another method known as the sale-price-to-appraisal ratio (SPAR). SPAR uses assessment information collected by valuers in some form or other to form a base period from which to calculate a change in overall price levels.

The simplest and most common method for generating an estimate of property values is of course a simple median measure of recent sales i.e. taking the median sale price in an area in a given year. Using the observation for a given year as the base period, price growth can be calculated as the percentage growth from the simple median in period to period . Because property markets are often segmented across the price range or ownership type (investment vs owner occupied) or geography etc., and because various macroeconomic conditions will often heavily influence the level of turnover in different segments of the market, the simple median index is often improved with various stratification techniques. Stratification attempts to control for the effect that exogenous conditions have on the turnover in different market segments and thereby controlling bias that might occur if only using the simple median. This works by splitting the market into several strata (for example 10 strata each comprising 10% of the total number of sales or splitting by property type) and taking the median of each. The index is then expressed as the mean of medians or similar. Stratified median indices should in theory be more readily comparable across the business cycle. While the simple and stratified median indices provide a straightforward estimate of house prices based on readily available data they suffer from sample bias and provide no information on property qualities.

Attempting to achieve a greater level of accuracy in their house price estimates, government and private sector institutions will often employ slightly more sophisticated techniques again for value estimation. The two most widely used HPI methodologies are the repeat sales estimator and hedonic regression. While both produce credible indices for the most part, the methods themselves suffer from significant drawbacks. In the simplest terms, the repeat sales estimator is potentially biased due to a sample selection problem: the distribution of houses that sell most frequently may differ from that of the population of houses. The hedonic regression model controls for variation in house characteristics statistically, and therefore suffers from model misspecification or unobserved variables, such as house quality. These shortcomings are reasonably well documented although both continue to be used widely.

These problems have recently led the RBNZ to investigate more accurate methods of constructing house price indices on behalf of the Real Estate Institute of New Zealand (REINZ). In March of this year the RBNZ benchmarked the SPAR method against the simple median and stratified median currently used by the REINZ. The SPAR methodology has been adopted by a few countries including Denmark, Sweden and the Netherlands. SPAR is the simple mean of the ratio of the most recent sale price to the most recent appraisal value. The mechanism by which it bypasses the model misspecification problem is that the appraisal value should incorporate more accurate information about the difficult-to-quantify aspects of a given property. The theory is that two otherwise identical properties that differ only in quality would have the same sale-price-to-appraisal ratio over the same period of time i.e. the true quality-adjusted level of price growth across the market. Across a large number of properties an appraisal value that accounts for quality differences will strip out any price growth effect attributable to quality when calculating an index. For example, if there are a large number of high quality properties sold in a given period, SPAR will adjust for what would otherwise be sample selection bias.

The fourth HPI methodology is the house price matching estimator. The matching estimator is a recent development. It has been proposed by McMillen in 2012. It is similar to the repeat sales method. Where repeat sales limits the data set to houses that have sold twice in the period of interest, the matching estimator pairs observationally similar houses together to perform the same calculation thereby widening the sample to include houses that have sold only once. The paper made some cursory suggestions as to the pre-processing methods by which one might determine matches but has left a considerable amount of room to contribute to the literature in that area. While not, to the best of my knowledge, currently used by any housing authority, it does possesses a number of desirable characteristics that in my opinion make it superior even to the SPAR HPI. I believe that it has a greater level of robustness to bias problems in HPIs currently in use. The construction of the index is relatively straightforward and relies only on data that is readily available. It is also well suited to specifically track indices across the house price distribution.

By far the most valuable property however, is that it promises better performance in small samples. This makes it particularly valuable for academic applications or measuring any number of policy or natural effects over a small geographic area. Locational amenities (or disamenities) may have sharp and idiosyncratic gradients and their effects may begin to be better understood only with a tool that in the simplest terms allows us to 'zoom in' to a greater degree. Examples might include sunshine hours that might differ greatly for houses on different sides of a street or on different sides of a hill. It would also include things like gang-occupied housing that, anecdotally at least, have an extremely localised effect. A matching estimator allows for greater 'power' to measure these effects.

McMillen also leaves room to investigate shortcomings of this index method. Specifically the problem of non-independence in latent variables. This might manifest as endogenous gentrification where improvements in quality (renovations) are not independent within a geographic area. High-income people are attracted to other high-income people. The unobserved income variable here is an enabler of renovation activity which would drive up the probability of further renovation activity. The SPAR index would appear to have an advantage here as the appraisals might reflect the general rate of renovation activity in the neighbourhood.

Despite the desirable properties of the recently adopted SPAR I believe that some form of matching estimator is actually a better solution to the house price index challenge. While it’s true that SPAR represents an improvement on the sample bias of the median and repeat sales methods and avoids altogether the model misspecification problem in the hedonic price estimator, it also suffers from its own problems. Recently Shi (2017) tested for stability in the SPAR index and found reason to be cautious regarding the reliability of estimates specifically with respect to both valuation assessment errors and temporal aggregation effect. Basically SPAR assumes that the distribution of quality in sales is stable over time. For instance, if there is an exogenous variable that has an influence on the distribution of sales over time (e.g., higher quality homes are selling more frequently in each successive period) then the ratio of sale-price-to-appraisal-value in high quality homes is over-represented in the index. This is effectively the same issue faced by the repeat sales estimator.

It's important to note that this first potential issue is predicated on the second. Systematic over-representation of the sale price ratio of a particular segment of the quality distribution is not a problem as long as the appraisal value across the quality distribution matches the market assessment of quality. That is, as long as the assessor correctly determines the market value of the property at the time of appraisal then the distributions will match and the sale price to appraisal ratio will be consistent across the quality distribution. If however, there is a systematic bias in valuations across the quality distribution then the over or under-representation of a SPAR for a particular portion of the quality distribution will cause the same sample selection bias that we're trying to move away from. The point is that there may be reasons for a systematic mismatch between the appraisal and market distributions.

Most importantly though, SPAR relies heavily on appraisal data. A cursory investigation into the appraisal process while writing this proposal indicates that these are neither complete nor accurate and may suffer from their own biases.

I propose to develop a matching estimator that leverages existing data efficiently to produce an accurate and robust housing index that can be relied upon even in small geographic areas or to calculate price growth across short timeframes. I hope to show that matching estimators can outperform existing measures and can be more reliably applied to a greater range of situations.

Urban price gradients: A comparison of pre and post-earthquake commercial centres and the effect on residential price gradients.

The traditional monocentric city model provides a theory for the organisation of land use and price based on the principle that residents or businesses are willing to pay a price for land up to the profit they can make from it. The key conclusion from the monocentric city model is that it predicts a negative exponential price gradient extending from the centre of the city. This central result broadly holds across a number of theoretical and empirical studies.

A large number of extensions have been made to the general monocentric city model of varying degrees of complexity. Different transport systems, industrial beltways, and agent preferences have been added to the original theory. Of particular relevance is the multi-centric city model. It was originally developed to describe the organisation of price surfaces with secondary commercial or industrial hubs arising because of natural geographic features or extreme size. While these theories have been validated empirically in cross-sections it is extremely rare to be presented with a set of circumstances that has the potential to cause a rapid shift in the underlying determinants of those price gradients. Observing how these gradients change under such circumstances moves the literature toward a better understanding of how catalysing elements of price gradient development such as land productivity inform, and in turn are informed by, urban planning and development of infrastructure.

The 2010 and 2011 earthquake events in Christchurch and resulting destruction of the city's CBD constitute, in purely economic terms, an interesting natural experiment. I propose to compare pre and post-earthquake cross-sections of residential property in Christchurch to determine the degree to which residential price gradients respond to the relocation of commercial centres. The purpose of this analysis is to observe the organic response of residential housing to planned versus improvised commercial centres. By observing the changes in the price gradients before and after the quake we can draw some conclusions about how planned and 'organically' occurring commercial centres affect price gradients. For example, how strongly do people respond to the relative location of jobs and to what degree are house prices determined by ‘stationary’ amenity values?

A pre-earthquake cross section of residential property values provides a benchmark price gradient based on a planned urban centre with appropriate infrastructure and zoning typical of small to medium scale, developed cities. Earthquakes in 2010 and 2011 which removed the CBD as the main locational amenity to which a residential price gradient was derived (in terms of the monocentric city model) might be expected to re-equilibrate around new commercial and industrial centres. These would form as businesses and hence jobs relocated themselves to areas that were not destroyed. I expect that this effect would be largely independent of other locational amenities arising from the natural geography (apart from the obvious change in landscape due to earthquake risk). Using techniques drawn from the multi-centric city literature I would likely carry out a series of estimates of price surfaces rather than simple price gradient estimates. One-dimensional price gradients might be inappropriate as house prices do not always move strictly monotonically with respect to distance. For example suburbs on the north side of the Christchurch city centre are more ‘posh’ than those located immediately to the south despite similar ‘as the crow flies’ distances from the CBD. Also, we might expect multiple new commercial centres in the absence of another single area with a high enough carrying capacity to replace the city centre.

These data provide a unique opportunity to perform an empirical study on organic price gradient formation. The degree to which house prices respond to the shift in commercial centres as businesses relocated away from the CBD will reveal information about the parameters of the price gradient models first developed in the monocentric city model. By making use of this natural experiment this work would make a contribution towards the ability to estimate indifference curves for the substitutability of housing services to consumption with respect to land productivity.

There are of course several issues that complicate this task. The first is price effects of earthquake related risk factors that became better known or at least, better publicised, after the earthquake. After the major earthquake events properties were assessed and placed on a risk spectrum for earthquake damage. Given there was greater transparency around these risk factors post-quake and the fact that what is traditionally a small probability event was at the forefront of residents minds means that these risk factors potentially have an impact on price that appears alongside the re-equilibration of the price gradients with respect to commercial centres. The second problem is that the effects of re-equilibrating price gradients, if large enough, could alter neighbourhood-level amenities not tied to the city's natural geography. A previously low income suburb for instance, might find that a shift in the residential price gradient triggers gentrification which would exaggerate the resulting price gradient estimates. This example illustrates that obviously not all locational amenities are independent of the productivity derived price gradient and would have to be accounted for in any investigation of price gradient changes.

Of particular value is that an alternative set of parameter estimates with respect to price gradient theory is that it may allow the measurement of the impact of urban planning and policy on land values and land use.

I already have access to a partial dataset of residential sales in Christchurch that has been used previously to investigate the direct impact of earthquake risk. Additional data are available through CoreLogic database.

Matching estimators in the construction of small sample house price indices: A comparison of studies of earthquake risk sensitivity in house prices.

This final chapter proposes using matching estimator techniques developed in the first chapter to address the problem of controlling for earthquake risk factor impact on price gradients as it applies to the analysis in the second chapter.

One of the crucial advantages of the matching estimator that I propose to examine in the first chapter is its performance in small samples. Here, I suggest a test case application based on the Christchurch dataset used in the previous chapter. In 2012, Geordie Reid carried out an investigation of house price response to earthquake risk factors in Christchurch. This is the direct impact of the earthquake risk ‘zoning’ that was carried out and subsequently published after the major earthquake events. He investigated to see if there was a statistically significant discount to housing based on risk 'zones' of differing intensities and found that there was no measurable impact of those risk assessments in areas that were still habitable.

I propose to revisit the analysis using the matching estimator rather than the hedonic regression methods in Reid. I expect that the improved small sample performance of the matching estimator will give me the ability to more accurately measure the effects of earthquake risk on residential property prices. This will be important to the proposed analysis on price gradient changes as outlined above.

By eliminating the model misspecification problems that may have occurred in the hedonic regression estimates of the initial investigation and increasing the ‘repeat sales’ sample size for a relatively small geographic area I hope to find that the impact of the publication of ‘technical categories’ on house prices to be significant.

The application of the matching estimator here will also be valuable to the initial HPI comparison chapter as a test case of the matching indicator performance across a real dataset. It would be reasonably trivial to carry out the analysis using the other methods as well to check the small sample performance and the subsequent impact on the findings.

My hope ultimately is that development of a reliable, relatively simple, and systematic approach to the application of the matching estimator along with initial empirical testing might contribute to its consideration as an alternative to current HPI methods in NZ.