A COMPARISON OF RESIDENTIAL RENTAL INDICES

BOB HARGREAVES and MIKE CHEN
Department of Finance, Banking and Property
Massey University, New Zealand

ABSTRACT

This paper compares and evaluates several different methods of measuring changes in residential rental levels for North Shore City in New Zealand. The indices compared were median, hedonic, repeat rent, weighted repeat rent and hybrid.

Technically the hybrid method appeared to be the most appropriate, but the trade off was this method is relatively costly in terms of operator time and data requirements. The repeat rent method is reasonably simple and has lower administrative costs. The median method is very easy to set up but suffers from both seasonal and constant quality difficulties. The hedonic method is also costly in terms of time and data requirements. On balance, the hybrid method was preferred.

Keywords: Rental indices, rental housing, measuring rents, New Zealand

INTRODUCTION

The accurate measurement of changes in residential rental levels is important for two main reasons. Firstly, both tenants and landlords frequently use existing contract rental figures as a benchmark in rent fixing procedures. They then update this amount by a percentage to reflect current market conditions. Secondly, changes in rental levels are an important component of the consumer price index. If these changes are not being measured accurately by Statistics New Zealand, then the Reserve Bank of New Zealand may use incorrect interest rate settings.

This paper uses residential rental data from North Shore City to compare and evaluate four main types of indices: median, hedonic, repeat and hybrid. For the purposes of this paper, the "best" index will accurately reflect changes in the rental market, be able to be easily updated and use data sources available from the public domain.

INDEX METHODS

Median

The traditional way of measuring changes in residential rental levels is to compare median rents over time by constructing a median chain index. This is done by selecting a specified time period as the base and setting the median rent in this period equal to a number, normally 100, as used in this paper. The median rent in the second and subsequent time periods is then calculated as a percentage change on the base period and the index is adjusted accordingly.

Data on private sector rents is available from the Ministry of Housing (Tenancy Bond Centre) in spreadsheet format. The median rent for any data set can be easily

calculated. The main disadvantage of using median rents is the constant quality change problem. That is, the type of property being transacted in one period may not be a representative sample of the total population of rental properties. This problem can arise from seasonal effects, such as tertiary students renting lower quality accommodation at the start of the academic year. Problems can also arise if properties being added to the rental housing stock are better quality than the average of properties within the existing rental pool. Similarly, upgrading the existing rental housing rather than simply maintaining it can lead to non-constant quality bias. Because the population of North Shore City was expanding at 2-3% per year and there was a substantial amount of new housing being built each year from 1992-1998, it is very likely that the constant quality problem is a real issue in this study.

Hedonic

The hedonic method involves the use of multiple regression analysis (MRA). MRA has been widely used as a tool for the mass appraisal of residential properties since the 1970's. To construct a MRA equation, the analyst collects information on transaction prices, transaction dates and quality attributes of properties in the data set.

Case, Pollakowksi and Wachter (1991) used the hedonic method to develop a price index in four American counties. Mark and Goldberg (1984) discussed alternative MRA approaches for the development of property price indices, and more recently Malpezzi, Chun and Green (1998) developed a house price index using log-linear form.

When using MRA methodology for the construction of a rental index, the analyst typically takes the amount of rent as the dependent variable and regresses this against a set of independent variables that describe the quality attributes of properties. By using time dummy variables, the drift of price movement over time from the constant term can be measured.

The accuracy of a hedonic index is dependent on how well the equation is specified. In practice, it is hard to specify a complete and appropriate set of property attributes because data on some attributes may be unavailable. For example, in this study, the street number of the rental properties was not available, so the rental data set could not be matched with the Quotable Value New Zealand database on individual property attributes. As public servants become more experienced in operating privacy legislation, it is likely database matching will be permitted provided that only aggregated results are published.

Repeat

Pioneering use of the repeat sales methodology was reported by Wyngarden (1927) and Wenzlick (1952). These early workers lacked the computation power of modern computers and relied on chain indices. Later Griliches (1961) applied hedonic methods to repeat sales for automobiles. The modern idea of a repeat sales index for property was developed by Bailey, Muth and Nourse (1963) and later refined and modified by Case and Shiller (1987, 1989 and 1990). This method also utilises MRA methodology, but avoids the difficulty of setting up the appropriate model and selecting the complete set of independent variables by basing the price index on repeat sales. The repeat method assumes that the price index of the underlying

properties in a geographical area can be represented by the price index found on a typical sub-group of properties (which transact twice or more during the same period). If the characteristics and quality of the sub-group properties remain unchanged, the price change between previous and current transactions is due to the time period that intervened.

The Bailey, Muth and Nourse (BMN) index is constructed by regressing the log price change between previous and current sales (dependent variable) on a set of time dummy variables (the only independent variables) for each property, using ordinary least squares regression.

The Case and Shiller refinement to repeat sales methodology challenged the assumption that the difference between the individual property log price and the citywide log price index is non-correlated through time. They showed that the variance is likely to increase with the time interval between sales and so is the regression residual. The reasons for these increases are firstly, the influence of depreciation over longer intervals, and secondly, the lengthening of the time interval making the previous rent setting less useful.

Accordingly, the Case Shiller method down-weights sales observations corresponding to larger time intervals between sales. This is done by a three step weighted regression where the first step follows the BMN method. The second step uses the residuals from step 1 as the dependent variable and regresses this on the holding time period. The third step repeats step 1 after the log price changes have been divided by the square root of the fitted value from step 2. In this paper, the Case Shiller method is referred to as the weighted repeat sales (WRS) method.

In this paper, both the BMN and WRS methods are tested on repeat rental data instead of repeat sales data. A common criticism of the repeat methods when applied to sales data is that a lot of data is wasted because houses only transact every 7-10 years and most of the sales in any given period are not repeat transactions. However, this criticism is not nearly as valid when applied to the rental market because the average length of a residential tenancy is less than two years. In 1996 there was a population of 12,367 private sector rental dwellings in North Shore City. During 1996, 6,685 new rental tenancies were recorded for the city.

Hybrid

The hybrid property sales index method was developed by Case and Quigley (1991), and further refined by Clapp et al (1991), Quigley (1995) and Eichholtz (1997). This approach combines elements of both the hedonic and repeat methods. The hybrid method normally uses all transaction data in the construction of the index and is less likely to be biased by the "starter home churn" problem. This means that first home sales, which tend to transact more than the market as a whole, are over-represented in some repeat sales indices. Sometimes these first homes are sold by the developer on a low deposit with a compensating higher price. When these homes sell under normal financing arrangements in the second hand market, the prices are reduced accordingly. This effect can put a downward bias on the index.

The hybrid model makes use of repeat transaction data to catch the time interactive effects on property price changes, while the hedonic part of the model deals with the

quality change problem and makes use of all transaction observations. For example, the Eichholtz study of buildings in Amsterdam used dummy variables to account for changes in building use over time as it switched from residential to commercial or vice versa.

The hybrid index method suffers from similar difficulties to the hedonic and repeat methods. The analyst still has to choose the appropriate functional form for the equation and needs to obtain a complete set of property attributes. Including all the transacting data may be more representative than just including repeat transactions, but there is still the question of how well the transacting data represents the total population.

The above constraints are, however, more valid for the house sales market than the rental market, since the turnover rate in the rental market is so much higher. Rental properties that had more than 10 tenancies over the 7 year period studied were excluded from this study on the grounds that they may have certain undesirable characteristics not reflected in the market in general.

DATA

Two sets of data were collected for this study. The primary data set contained information on the geographical location of the rentals, the dwelling type, the number of bedrooms, the rent paid, and the tenancy commencement date. The sample was taken from the entire urban area of North Shore City and only included those individual private sector residential tenancies registered in the Ministry of Housing (Tenancy Bond Centre) during the period of 1992 to 1998. This sample is thought to be a good representation of the private tenancy population, since landlords must lodge bond money with the Tenancy Bond Centre if a bond is collected. For reasons of privacy, the data set did not contain individual street addresses, but properties could be located to within a census mesh block.

Over 39,000 rental observations were recorded in the data set. From these, 35,431 effective tenancies were identified (including 9,256 single tenancies and 26,175 repeat tenancies). Observations that appeared to have data entry errors, or that may have biased the indices were excluded from the analysis; those included:

- Properties identified neither as flats nor as houses.
- Properties with area geographic identity codes not found for North Shore City on the map of Northern Auckland Urban Zone 1991.
- Properties with no bedroom or more than four bedrooms.
- Properties which had tenancies 10 times or more during the 7 years sample period.
- Rents out of the ranges of $70 to $350 for one-bedroom properties, $100 to $500 for two-bedroom properties, $130 to $700 for three-bedroom properties and $160 to $1,000 for four-bedroom properties. Although the dollar cut off amounts used here were somewhat arbitrary, they were

based on the authors' experience of what was considered reasonable.

- Repeat rents with no time interval between tenancies.
- Repeat rents with rent change out of the range of "-20% to +50%" between tenancies.

A summary of the rental data is contained in Table 1. The data set is dominated by three-bedroom houses, and repeat tenancies are much more common than single tenancies. A single tenancy is a property that was only rented once during the 7 year period studied. A repeat tenancy is a property renting between 2 and 9 times over the study period. Only 12% of the repeat rent data set had more than 5 tenancies.

Table 1: Summary of Effective Rents

Observation Type	Single Tenancies		Repeat Tenancies		All Tenancies Pooled	
	Rental Mean $	Observations	Rental Mean $	Observations	Rental Mean $	Observations
By no. of bedrooms:						
One-bedroom	163.26	770	168.45	2535	167.24	3305
Two-bedroom	224.60	2885	217.57	9429	219.22	12314
Three-bedroom	286.00	4560	279.45	12091	281.24	16651
Four-bedroom	355.06	1041	343.83	2120	347.53	3161
By dwelling type:						
Flats	217.35	3243	210.02	11156	211.67	14399
Houses	289.80	6013	282.53	15019	284.61	21032
By wards:						
East Coast Bays	272.90	1495	261.80	4242	264.69	5737
Glenfield	255.68	2186	244.82	6021	247.71	8207
Takapuna	276.10	2324	257.06	7503	261.57	9827
Onewa	250.46	2167	241.19	5690	243.74	7857
Devonport	273.17	1084	257.66	2719	262.08	3803
Summary	264.41	9256	251.62	26175	254.97	35431

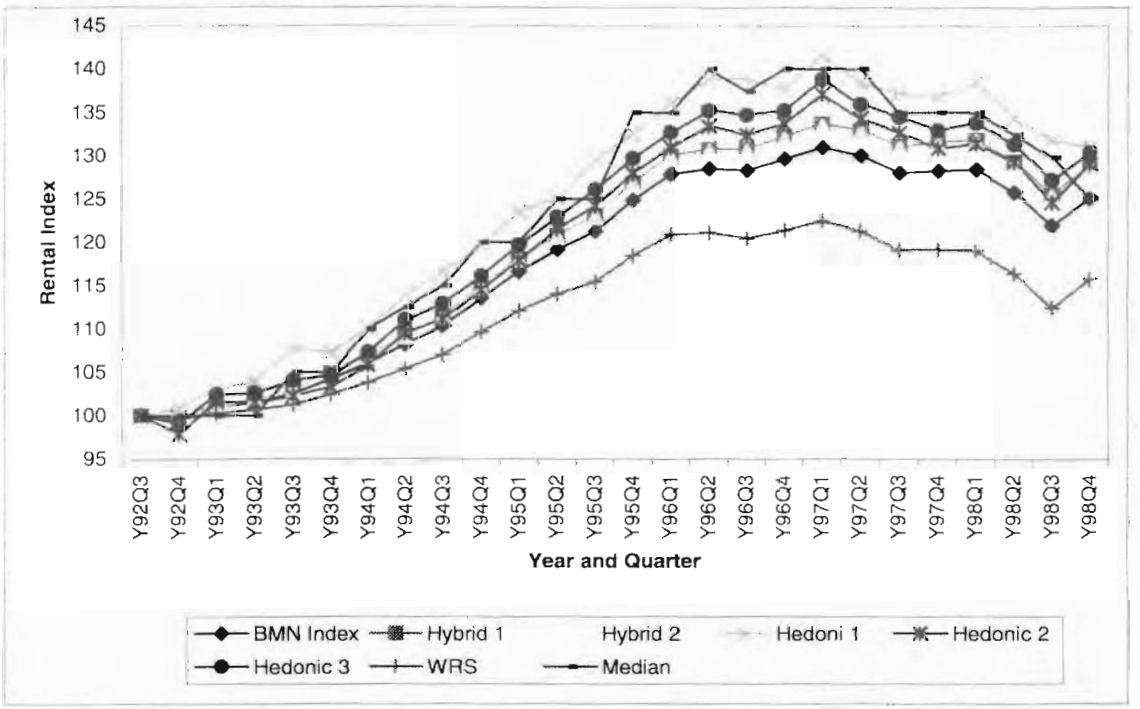
The secondary data set contains information on demographic changes and household incomes taken from the 5 yearly Census conducted by Statistics New Zealand (1991), (1996). Data on average house prices within the wards of North Shore City was obtained from Quotable Value New Zealand.

At the time of the 1996 census, North Shore City had a population of 171,494. The city is situated to the north of the Auckland harbour and is one of four cities comprising the Auckland region. Population growth averages 2-3% per year and approximately 20% of dwellings are private sector rentals. The rate of population growth is amongst the highest in the country. During the period 1993-1997, growth in private sector residential tenancies averaged 5% per year.

RESULTS AND ANALYSIS

This section of the paper discusses eight rental indices that the authors have developed for North Shore City. These indices are shown in Figure 1 below. There is one median index, three hedonic indices, two repeat sales indices and two hybrid indices. The first repeat index is based on the BMN model and the second utilises the WRS model.

Figure 1: Index Comparison



The first hedonic model uses single tenancy transactions; the second, repeat tenancies and the third, all tenancies. Both hybrid models use just repeat tenancies. Hybrid model 1 has a very similar functional form to the BMN repeat index but includes a new independent variable to measure changes in rents due to the time interval between tenancies. This is done by taking the log of the time interval between tenancies.

Hybrid model 2 includes additional independent variables as follows:

- Log of the number of bedrooms
- Dummy variables for location by wards
- Dummy variables for dwelling type (house or flat)
- Log of yearly median household income growth by wards
- Log of yearly population growth by wards
- Log of average house prices by wards

The results of the regressions are presented in Table 2.

Table 2: Regression Results Comparison

Types of Model	R-square	Standard Error of Estimate
Hedonic Model 1 (single tenancies, n=9,256)	0.624	0.1875
Hedonic Model 2 (repeat tenancies, n=26,175)	0.636	0.1728
Hedonic Model 3 (all data, n=35,431)	0.633	0.1772
BMN Model (repeat tenancies, n=26,175)	0.460	0.0708
Hybrid Model 1 (repeat tenancies, n=26,175)	0.460	0.0708
Hybrid Model 2 (repeat tenancies, n=26,175)	0.465	0.0705
WRS Model (repeat tenancies, n=26,175)	0.373	0.0709

It is interesting to note that hedonic model 1 appears to closely track the median index and is subject to relatively large fluctuations. This is probably because hedonic model 1 only uses single tenancies, and these suffer more from constant quality upward drag and seasonal fluctuations than the other hedonic models. All the hedonic models appear to explain around 63% of the variation in rents, with the most powerful explanatory variable being the number of bedrooms. However, the R-square statistic must be treated with caution as the standard error of estimate is relatively larger for the hedonic models than for models incorporating the repeat method.

Figure 1 also illustrates the volatility of the indices. During the period when rents were increasing from 1993 to the first quarter of 1997, the median index and the hedonic indices increased significantly faster and fluctuated more than the other indices. In the following time period, the gap is similar. This is likely to be due to the quality change problem as more high rent properties entered the market. The median index is not a good measure of rental movements, nor are, the hedonic indices in this study because there are not enough property attributes to deal with the quality change problem.

In this study (1992-1998), properties showing repeated renting are much more likely to be representative of the rental population than single tenancies. This is because a residential rental that lasts for more than five years may well be at less than market rent, as landlords are reluctant to increase the rent for sitting tenants.

Figure I shows that the hybrid methods and the repeat methods are less volatile than the hedonic methods and do not reach the same peak. Both the repeat rent methods are at the bottom end of the graph, particularly the WRS method (with only a 22% increase at its peak). The down-weighting factor applied by Case and Shiller (to paired transactions corresponding to longer time intervals between house sales) may not be appropriate in the faster churning rental market context. The BMN index is a little lower than the hybrid methods (134 versus 131 at its peak) and may be still underestimating actual movements in the rental market.

The two hybrid indices produce very similar results. This shows that the explanatory power of the additional variables included in hybrid model 2 are negligible. However, the log of the time interval between tenancies is a significant variable with a correlation of 29.9% to the dependent variable. The table of correlation coefficients is produced in Table 3. The explanation for the names of the abbreviated variables is as follows:

LN P*it*	Log of rental price, the dependent variable
R *itt'*	Change in the log rent between tenancies
LN_INTER	Log of time interval between tenancies
LN_POP	Log of annual population growth (1991-1996) by individual wards
LN_INC	Log of median household income growth (1986-1996) by individual wards
LN_SALES	Log of average house price (1998) by individual ward
LN_ROOMS	Log of number of bedrooms

Clearly, the number of bedrooms is a very important variable when explaining differences in rental levels between properties. However, this variable drops out when using the repeat methods. Details of the equations for the hedonic models are contained in Appendix 1 and the balance of the models are shown in Appendix 2. The models only use dummy variables for time and location by wards. It is acknowledged that the actual regression coefficients must be treated with caution, as generic ward data does not reflect individual property attributes.

Table 3: Correlation Coefficients

		LN P*it*	R *itt'*	LN_INTER	LN_POP	LN_INC	LN_SALES	LN_ROOMS
LN P*it*	Pearson Correlation	1.000	.151**	.086**	-.004	-.011*	.114**	**.656****
	Sig. (2-tailed)	.	.000	.000	.441	.046	.000	.000
	N	35431	17509	16616	35431	35431	35381	35431
R *itt'*	Pearson Correlation	.151**	1.000	**.299****	-.017*	-.008	.002	-.026**
	Sig. (2-tailed)	.000	.	.000	.026	.298	.784	.001
	N	17509	17509	16612	17509	17509	17489	17509
LN_INTER	Pearson Correlation	.086**	.299**	1.000	-.017*	-.020**	-.013	.039**
	Sig. (2-tailed)	.000	.000	.	.029	.009	.085	.000
	N	16616	16612	16616	16616	16616	16596	16616
LN_POP	Pearson Correlation	-.004	-.017*	-.017*	1.000	.651**	-.297**	.128**
	Sig. (2-tailed)	.441	.026	.029	.	.000	.000	.000
	N	35431	17509	16616	35431	35431	35381	35431
LN_INC	Pearson Correlation	-.011*	-.008	-.020**	.651**	1.000	-.245**	.063**
	Sig. (2-tailed)	.046	.298	.009	.000	.	.000	.000
	N	35431	17509	16616	35431	35431	35381	35431
LN_SALES	Pearson Correlation	.114*	.002	-.013	-.297**	-.245**	1.000	-.157**
	Sig. (2-tailed)	.000	.784	.085	.000	.000	.	.000
	N	35381	17489	16596	35381	35381	35381	35381
LN_ROOMS	Pearson Correlation	.656**	-.026**	.039**	.128**	.063**	-.157**	1.000
	Sig. (2-tailed)	.000	.001	.000	.000	.000	.000	.
	N	35431	17509	16616	35431	35431	35381	35431

** Correlation is significant at the 0.01 level (2-tailed).
* Correlation is significant at the 0.05 level (2-tailed).

SUMMARY AND CONCLUSIONS

The issue of which index is "best" is not clear-cut, because the answer depends on who is using the index, what they are using the index for and the costs and benefits of each option. Furthermore, there is no way of determining what the actual movement in rentals is, since no precise benchmark exists. The median index is inexpensive to compile and is easily updated, but it is likely to over-estimate actual changes, because

it may violate the constant quality requirement. In a city such as North Shore where there has been significant growth in the rental housing stock, the median method almost certainly does violate constant quality.

In theory, the hedonic methods outlined should work well, but there are considerable practical difficulties in being able to capture all the important independent variables. The data does exist in various databases, but only some of these are in the public domain. Further, matching individual property records from merged databases may violate privacy legislation in New Zealand.

The two repeat methods discussed in this paper were developed in relation to house price indices. Neither of these models is calibrated to work in the much faster churning residential rental market. The advantage of the repeat methods is their relatively low data requirement, thereby reducing the cost of assembling such indices. This disadvantage of repeat sales not being representative of the whole housing market is again largely overcome in the rental market (where the average length of tenancy is less than two years). In the absence of a property rental database recording upgrades to rental properties, the model assumes repairs and maintenance to property equals depreciation. This may not be a valid assumption.

The hybrid methods discussed in this paper combine elements of the repeat and hedonic methods. Hybrid model 1 achieves the same result as the hybrid model 2 which is more complicated and has more extensive data requirements. Hybrid model 1 follows the same methodology as the BMN repeat method, but has an additional variable to account for the time interval between tenancies.

In conclusion, the authors recommend hybrid model 1 as the model to use. The data can be obtained from one source, and the model is less complicated than the hedonic models and hybrid model 2. In addition, hybrid model 1 appears to be accurate, and has an acceptable level of volatility.

REFERENCES

Bailey, M.J., Muth, R.F., & Nourse, H.O. (1963). A Regression Method for Real Estate Price Index Construction. Journal of the American Statistical Association, 58, 933-942.

Case, B., & Quigley, J.M. (1991). The Dynamics of Real Estate Prices. Review of Economics and Statistics, 73(3),50-58.

Case, B., Pollakowski, H.O., & Wachter, S.M. (1991). On Choosing Among House Price Index Methodologies. Journal of the American Real Estate and Urban Economics Association, 19(3), 286-307.

Case, K.E., & Shiller, R.E. (1989). The Efficiency of the Market for Single-Family Homes. The American Economic Review, 79(1), 125-137.

Case, K.E., & Shiller, R.J. (1987). Prices of Single Family Homes Since 1970: New Indexes for Four Cities. New England Economic Review, September/October, 45-56.

Case, K.E., & Shiller, R.J. (1990). Forecasting Prices and Excess Returns in the Housing Market. Journal of the American Real Estate and Urban Economics Association, 18(3), 253-273.

Clapp, J.M.., Giaccotto, C., & Tirtiroglu, D. (1991). Housing Price Indices Based on All Transactions Compared to Repeat Subsamples. Journal of the American Real Estate and Urban Economics Association, 19(3), 270-285.

Eichholtz, P.M.A. (1997). A Long Run House Price Index: The Herengracht Index, 1628-1973. Journal of the American Real Estate and Urban Economics Association, 25(2), 175-192.

Griliches, Z. (1961). Hedonic Price Indices for Automobiles: An Econometric Analysis of Quality Change. Washington: Government Printing Office, p. 173-196.

Malpezzi, S., Chun, G.H. & Green, R.K. (1998). New Place-to-Place Housing Price Indices for U.S. Metropolitan Areas, and Their Determinants. Journal of the American Real Estate and Urban Economics Association, 26(2), 235-274.

Mark, J.H. & Goldberg, M.A. (1984). Alternative Housing Price Indices: An Evaluation. Journal of the American Real Estate and Urban Economics Association, 12(1), 30-49.

Palmquist, R.B. (1982). Measuring Environmental Effects on Property Values without Hedonic Regressions. Journal of Urban Economics, 11(3), 333-47.

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

Statistics New Zealand. (1991). 1991 Census.
Statistics New Zealand. (1996). 1996 Census.

Wenzlick, T. (1952, December 24). As I see the fluctuations in the selling prices of single-family residences. The Real Estate Analyst, 541-548.

Wyngarden, H. (1927). An Index of Local Real Estate Prices. Michigan Business Studies, 1(2),15-22.

Appendix 1: Details of Hedonic Models 1, 2 & 3*

Variables	Regression Coefficients		
	Model 1	Model 2	Model 3
(Constant)	1.971	2.988	2.698
Y92Q3	.089	- .012	.018
Y92Q4	.098	- .034	.009
Y93Q1	.120	.003	.042
Y93Q2	.128	.002	.044
Y93Q3	.164	.010	.058
Y93Q4	.159	.020	.064
Y94Q1	.186	.044	.088
Y94Q2	.218	.078	.123
Y94Q3	.243	.093	.140
Y94Q4	.269	.123	.168
Y95Q1	.300	.152	.198
Y95Q2	.313	.183	.225
Y95Q3	.348	.205	.250
Y95Q4	.371	.235	.279
Y96Q1	.398	.258	.302
Y96Q2	.418	.277	.321
Y96Q3	.418	.269	.317
Y96Q4	.410	.279	.320
Y97Q1	.437	.304	.346
Y97Q2	.415	.283	.326
Y97Q3	.406	.271	.315
Y97Q4	.403	.257	.304
Y98Q1	.415	.261	.310
Y98Q2	.384	.245	.291
Y98Q3	.366	.208	.259
Y98Q4	.360	.243	.284
GLENFIELD	.046	.058	.056
TAKAPUNA	.038	.039	.039
ONEWA	- .0004	- .009	- .059
LN_ROOMS	.524	.460	.478
HOUSES	.073	.105	.096
LN_SALES	.234	.178	.194
LN_POP	- .0378	- .059	- .054

*: Dependent variable is LN_PIT.

Appendix 2: Details of Hybrid Models, BMN Model and WRS Model

Variables	Regression Coefficients			
	Hybrid 1	Hybrid 2	BMN	WRS
(Constant)	- .006	.082	- .008	.001
Y92Q3	.011	.011	.009	- .003
Y92Q4	.008	.009	.005	.002
Y93Q1	.022	.021	.020	.006
Y93Q2	.029	.030	.023	.012
Y93Q3	.041	.041	.034	.024
Y93Q4	.058	.059	.050	.037
Y94Q1	.077	.077	.068	.052
Y94Q2	.096	.098	.088	.068
Y94Q3	.119	.120	.108	.092
Y94Q4	.149	.149	.136	.114
Y95Q1	.177	.177	.163	.131
Y95Q2	.200	.199	.184	.131
Y95Q3	.218	.218	.202	.144
Y95Q4	.249	.248	.231	.170
Y96Q1	.274	.273	.255	.190
Y96Q2	.280	.279	.260	.191
Y96Q3	.279	.279	.259	.186
Y96Q4	.291	.290	.269	.194
Y97Q1	.302	.301	.279	.203
Y97Q2	.296	.296	.272	.193
Y97Q3	.282	.282	.256	.175
Y97Q4	.285	.285	.258	.175
Y98Q1	.288	.287	.259	.174
Y98Q2	.269	.268	.239	.153
Y98Q3	.238	.237	.208	.117
Y98Q4	.265	.264	.233	.147
LN_INTER	- .006	- .005		
LN_SALES		- .005		
GLENFIELD		- .001		
TAKAPUNA		- .001		
ONEWA		- .0003		
DEVONPORT		.0002		
LN_ROOMS		- .011		
HOUSES		- .006		

For hybrid 1, hybrid 2 and BMN, dependent variable is R_ITT

For WRS, dependent variable is WEIGHTED.