

# The Quality of Data and Data Availability for Property Research

Ge, X J and Harfield, T  
School of the Built Environment, Unitec Institute of Technology, New Zealand

## Abstract

Property market performances have great impact on the economy as a whole. Property research is one of the disciplines of scientific research that provide significant information for decision-makers. However, scientific research requires quality data. Quality data plays a vital role in providing reliable and valid information for property market performance. This paper discusses the state of property data in terms of both quality and availability. A possible solution is made concerning the current lack of adequate availability of quality historical datasets for property researcher.

Keywords: Quality of data, Data availability, Property market.

## Introduction

Property markets can affect price stability and distort resource allocation. Evidences suggest that property markets play an important role to the macro economy. The construction and property industry in Hong Kong has contributed over 20 percent of GDP since 1982 (Walker, *et al.*, 1995). Property and construction company stocks contributed 25 percent of Hong Kong's stock market capitalization, and over 60 percent (on average) of capital investment expenditure have been in property since 1983 (Newell & Chau, 1996). More than 45 percent of all bank loans are directly tied to properties (Chan, *et al.*, 2001). In Australia, dwelling sales have expanded from \$90 billion or 13 per cent of GDP in 1999 to \$156 billion or 21 per cent of GDP in 2003 (Ballard, 2004). It was recorded that the residential property sector contributes around 17% of the total GDP in Auckland in 2005 (New Zealand Government, 2006). Total value of owner occupied domestic stock was estimated to be £ 1450 billion at UK in 2000 (Jenkins, 2000). This suggests that researches in property market performance, forecasting supply and demand in the market place and the impact on economy are important. Quality researches may not only assist government to formulate macro and micro economic policies, but also for investors to understand the market and making investment decisions.

Data quality and availability may become even more important and thus enhance the quality researches.

Predicting property market movements is one application of economic forecasting, which is an attempt to develop an educated view of the future for decision-making purposes). One definition of economic forecasting is an analysis of economic facts which takes into account theories such as the business cycle and supply-demand trends. However, analysis is only possible because specific data (economic facts) are available to the analyst (Carnot *et al*, 2005)

The balance of this paper focuses on issues concerning data which is used for property market analysis. The paper will first discuss the nature of data and the criteria of quality in a number of countries. Topics concerning property data including the availability of macro economic data and the link of that data to property markets will be discussed. This will be followed by an exploration of the quality and accessibility of some property indices. The last section will provide a new model of availability of data which will be beneficial to all property researchers.

## **The Quality of data**

### *Nature of data*

Forecasters use a common set of tools. The tools consist of data sets and methodologies which are used to create the datasets. Data are the variables or factors which are used to measure the development of economies. Analysis of the data is usually part of a set of statistical standards within as a coherent framework. Although the framework may also be standard, the data may be interpreted from a variety of perspectives. Thus, forecasting is considered by some to be more of an art than a science (Carnot *et al*, 2005).

Because of the importance of economic forecasting to the economy, many governments have taken on the task of collecting data. It is assumed that with government control of the collection of data, data will be both reliable and valid. The features of reliability are dependability, stability, consistency, predictability, and accuracy (Burns, 2000). However,

data is often incomplete or inaccurate at the time of collection. For example, population figures collected one census day need to be up-dated at regular intervals between each five year census. This process of correction is undertaken to ensure that the data is considered to have validity which is the requirement that data measure what they say they measure. So even though interpretation of data changes over time changes, the cornerstone of good forecasting the reliability and validity of collected data (Cook, 2004).

### ***Criteria for quality***

That is not to say that just because governments collect data it is perceived as reliable and valid. Data sets need also to have some historical validity to gain acceptance in the international forecasting community. If data are produced over time, with well defined explanations for changes in methodology and variable measurement, then data sets will be considered consistent. For example, many countries that create national accounts focus on a set of productive industries which were categorised in the 18<sup>th</sup> century when agriculture was the main provider of national wealth. The service sector, which in the 21<sup>st</sup> century makes up the majority of many economies, has had to be included into the data collection of the productive endeavour. Methodologies to measure the contribution of the service sector have been developed with valid statistical justifications to provide historical data sets (Dixon & Hamilton, 1996).

A set of criteria which indicates high quality of data is needed because economic endeavour changes over time. The first criterion is that the data be used by reputable organisations not related to the data collectors. For example, employment data collected by Statistics New Zealand is used by multinational banks to forecast the business climate. The second criterion is the regular output of data to provide an historical data set which takes into account both statistical and definitional revision such as the GPD. The third criterion is that the data collectors have an international reputation, all of the OCED countries regularly use government collected data (Blades & Roberts, 2002).

While the collection of national data is never comprehensive, national governments remain the most likely organizations to collect aggregate data. However, forecasters who focus on property performance, use macroeconomic data as well as microeconomic data that are specifically related to property.

### *Criteria for property data*

The primary aim in collecting data is to maximise the amount and accuracy of transfer of meaning from researcher (Fellows, 2003). The criteria for quality property data can be measured by the use of publicly available raw data, the use of internationally accepted factors and comprehensiveness of raw data.

Most publicly available raw data are facts. For example, the data are recorded in the Land Registry is the legal right information, that available for publicly accessing for a fee. The data generally includes ownership and title, land-related information, property size and structure, year of built, and used material. The sales price of properties and transaction data from real estate agents is also raw data. The use of such publicly available data to create proprietary indices suggests that the proprietary indices are considered to have statistical validity.

Internationally accepted factors for the indices are indicators of both reliability and validity of datasets. For example, the housing price index and the rental index published in the *Hong Kong Property Review* by the Census and Statistics Department. Property researchers rely such data which specifically relates to local standards that have gained international acceptance. The building consents, which provides data by local government on both the number of square metres as well as the projected expenditure for any permitted project, is an example of data collected in relation to internationally accepted building standards. Both government and private organisations can data which is internationally recognised. Some private organisations such as banks and consultants construct indicis and data for their own usage. These data sets may to be used by worldwide investors, analyst and potential customers. The obvious examples are the IPD Property Index provided by the Investment Property Databank and Russell Property Index which is used in Australia and Canada.

Comprehensiveness of raw data is an important component of reliability. Property has a combination of special characteristics which are different from other commodities; four specific categories of features as noted in table 1:

Table 1 Special Characteristics of Property

Physical	Legal	Market	Costs
durability	variable rights		high purchase price
spatial immobile	informational asymmetries	thinness of the market	Transaction costs

heterogeneity		involvement of governments in housing	
multi-dimensional		related input markets	

Because of the four interrelated categories, there are three necessary elements for property forecasting data. The first requirement is a very large sample size which provides a good representation of the population. The second requirement is a widest range of market variables. Information on different types of property is never sufficient. The third requirement is data reliability which means consistency with accuracy (Burns, 2000). So even though interpretation of data changes over time, transparency is expected in historical datasets to incorporate the changes of recording and reasons for those changes.

*Accuracy of property data*

Lum (2004) suggested that the accurate measurement of real estate price movements is an issue of great concern to both market participants and policymakers who rely on price signals for decision-makings. Accuracy is a problem because of the fundamental heterogeneity of all property market data. For example, the unique location every building occupies (Dunse, *et al.*, 1998), the physical design, neighbourhood qualities, and such factors as convenience as measured by the distance to the city centre or public services (Omar & Ruddock, 2002). These features create measurement problems which has not be solved to-date.

There are serious defects in the way property indices are computed in many Commonwealth countries according to Lum (2004). An example is a methodological issue that concerns the choice of weighting procedures and the index number formula. Another problem found in measuring the capital value is the large degree of variation in the weightings to create the indexes. IPD Property index adopts capital value weighted, appraisal-based, chain-linked and time-weighted accumulation index. This practice causes inaccuracy in the attempted analysis.

Another methodological issue is the interpretation of the factors. In the case of property index, the Ratings and Valuation Department in Hong Kong apply transaction-based index by using government valuation to form value factors; whereas Quotable Value New Zealand applies transaction-based index by using government valuation to form price ratios.

In addition, when differences in composition, construction and disaggregation are encountered measurement issues arise. The composition of the published measures varies

significantly by sample size, type and size of fund, distribution by sector and region. In term of data construction, significant differences exist in the way returns are calculated and methodology. Some annual returns are calculated on a money-weighted basis, some on a time-weighted basis depending on the availability of data. When using disaggregation, most measures provide a breakdown of composition and results by sector according to Morrell (1995) are not particularly helpful.

### ***Completeness of property data***

Rental forecasting is an important element of the property investment process. Most rental forecasts are derived from econometric models explicitly or implicitly based on economic relationships (Mitchell & McNamara, 1997). The forecasting requires historical time series data.

However, obtaining time series market data are often difficult for property researchers. When data are available they are often incomplete. For example, house price index and rental index are often not accessible or incomplete. Data from building society surveys, market reports, auction results and information published in the press are often too general, out of date or are published too infrequently. (Wyatt, 1996)

Dunse, *et al.* (1998) discussed the completeness of published property data sources and the extent to which confidence can be place on research. They claim that the statistical information on the commercial property market provided by the government in England and Wales is very limited. The time series is not produced at sufficiently regular intervals for property investors to make informed choices. Property researchers and economic forecasters in general agree that datasets are incomplete if they lack an historical dimension.

Property research requires to access to comprehensive, reliable and timely evidence of property transactions in order to make informed predictions. The suggestion by Adair *et al* (1995) that performance of the commercial property market within Belfast was influenced by a combination of macro-and micro-economic indicators. For these researchers the macroeconomic inputs such as GDP and CPI along with local economic conditions relevant to the local property market including grant regimes are all factors which must be addressed in a quality analysis.

*Data required for property modelling*

Much of the property research focuses on valuation. A number of scholars suggest that both quality macro and micro data are required to develop property valuation models. Tse and Love (2000) applied hedonic price model for residential property valuation in Hong Kong. Property price, floor area, age, availability of car park, availability of a shopping centre, availability of sport facilities, cemetery view and estate type of housing units are the variables used for the valuation. The variables used by Wyatt (1999) also include factors such as social, economic, planning and environmental attributes. The variables most frequently used for property research has been identified in table 2 which indicates a mix of macro and micro elements.

Table 2 Principal Determinants for Net Effective Rent Forecasts

Property market determinants	Office	*	Industrial	*	Retail shops	*
Property	Historical rents	2	Historical rents Construction activity Construction costs	3	Historical rent Construction activity Occupancy costs	2 4 1
	Construction activity	1				
	Construction costs	0				
	Vacancy rates	1				
	Workspace ratio	8				
	Lease incentives	5 1				
Space	GDP	4	GDP Employed persons Estimated utilisation capacity	10 3 1	GDP Employment Retail turnover Population Private disposable income Consumer sentiment index	2 1 7 2 1 4
	Office employment	9				
	ASX – Indices	1				
	Job adverts	1				
Capital	Interest rates	1	Interest rates Inflation	2 2	Inflation	1

\* Selections by Property organizations. (Source: Higgins, 2000, p. 64)

A reduced-form property price model has been traditional used for residential property (Reichert, 1990). The determinants of quantity demand for housing considered demographic factors, housing-related elements such as unit transaction volume, and macroeconomic variables, such as GDP, CPI, etc. The supply of housing is a function of housing prices, construction costs, building consent, interest rates, material costs and labour costs, and land supply. Historical time serial data were required for modelling, but it is not always available.

**Availability of property data**

As noted above the property researcher has to take care in finding data that are both reliable and valid. Even if data are considered to have both these qualities, access to the data may be

limited. This section will outline some problems concerning data availability as well as a suggesting solution to the problem.

### *Sources of property data*

Property researchers can find some datasets which meet the quality critical mentioned in section one. These datasets are available from both government and private organisations in either print or electronic form. Table 3 shows local property index. the sources of property indices for the selected country.

Table 3: Property Index: International Sources

Country	Index	Provided by
Australia	BOMA-Russell Property Index (Established) House Price Indices: Eight Capital Cities Property Council of Australia Investment Performance Index	Building Owners and Managers Association Australian Bureau of Statistics  Property Council of Australia
Canada	New Housing Price Index Russell-Canadian Property Index	Statistics Canada Frank Russell Canada Limited
Hong Kong	Property Price Index Centa-city Index	Ratings and valuation department Centaline Property Agency Limited
Malaysia	Malaysian House Price Index	Ministry of Finance, Malaysia
New Zealand	QVNZ House Price Index Property Council of NZ Investment Performance Index	Quotable Value New Zealand Limited Property Council of New Zealand
Singapore	URA Property Price Index	Urban Redevelopment Authority
UK	IPD Property Index Nationwide House Price Index Halifax House Price Index	Investment Property Databank Nationwide Building Society Bank of Halifax

(Source: Lum, 2004, pp. 26)

Brown and Matysiak (1995) mentioned that the history of reliable commercial property indices in the UK is relatively short. Jones Lang Wootton (JLW) and Investment Property Databank (IPD) are the main property indices to the research of property performance, return and valuation in UK. However, Morrell (1995) suggests that the creation of a property index is problematic by comparing the availability of raw data for the equity investment markets and the property investment market as suggested in table 4.

Property Indexes providers do not have a consistent policy regarding availability. In New Zealand valuation data was a service available through a government agency until the agency was privatised in the early 1990s. Quotable Value New Zealand (QVNZ) retains the rights to

the use of historical data, which is required by law, with other private providers such as the Property Council of New Zealand and the Real Estate Institute of New Zealand (REINZ).

In Hong Kong the government has also collected property data from at least the 1980s. In 1999, Centa-City Index (CCI) and Centa-City Leading Index (CCL) were launched by City University of Hong Kong (CityU). The CCI is based on all transaction records registered at the land Registry of the HKSAR Government. The CCL is a weekly index based on preliminary contract prices in centaline transactions, to ascertain a more precise picture of the market (Leung, 2006). The indices provide information and trend for potential home-buyers and sellers.

The internationalisation of the property market has pushed organisations to provide analysis of purchase behaviour. These indexes, usually not available prior to the mid 1990s, are now found in a number of localities. For example, Hong Kong Polytechnic University (PolyU) has produced BRE Confidence Index for Residential property (BRE index) (Wong, 2006). In New Zealand, the ASB Bank began producing this type of Index in the late 1990s. Although new indexes are being produced more frequently than in the past, the problems concerning the availability of the **data** that makes up the indexes as discussed by Morrell (1995) remains problematic.

#### *Problems of availability of property data*

While it is desirable to provide investors with credible analysis based on reliable data, analysis is not sufficient for property researchers, they require the raw data. The type, extent, and volume of data that are available for those who analyse equity markets, is all but absent for the forecaster in the property market. Table 4 compares the favourable situation of equity market data to that less than perfect situation for the property market.

Table 4: Comparison Availability of Data for Equity and Property Market

Equity market data	Property Market data
traded prices and dividends is a public process	Data concerning property sales and rental is a private process
Data are easily aggregated	Data are not easily integrated
variables use is standardised	No agreement on what constitutes a variable
Analysis methodology is standardised	Wide variety of methodologies
Data are freely available	Many data sets require the outlay of cash
data produced rapidly	Significant delays in assembling data

There are three issues which relate to availability of data. The first problem is that data are hard to obtain. The private nature of property transactions is traditional values of confidentiality constraints and conservative attitudes. In addition, obtaining time series market data is often difficult for property researchers. For example, house price index and rental index are often not accessible due to the proprietary nature of many of the indices as noted above. In addition, some available information from banks or real estate agents may not be able to accessible to researchers because of legislative restrictions on data release to the public (Wyatt, 1996).

In addition some data are not freely available to the general public or to property researchers. Cost can be incurred for historical data or some micro-data. In New Zealand micro-data such as property types, construction materials, property features, and environmental condition is only available if purchased, which could be great if a time series is required.

The second barrier to access of data for research purposes is that the data that are available are not necessarily the data that researchers want. Wyatt (1996) studied the property valuation using a geographical information system. He required data for property transactions to be able to predict future property values. The lack of regional data for the property markets of England and Wales prevented him of completing his task. Other authors also remark on the problems the difficulties of the research process due to lack of large-scale micro-datasets (Leishman & Watkins, 2002).

The third pitfall for property researchers is the problems with dataset format. Computers, now the most common research tool, have the ability to use electronic datasets. The government agencies that collect aggregate data, such as census or productivity, have added electronic datasets to the traditional printed raw data. However, the availability of electronic datasets has met a demand, but also has led to an expectation that all datasets will be available in electronic format. For example, the New Zealand birth and death rates are available back into the 1872, but the migration data is only available in electronic form from 1996.

#### *Solution to problem of lack of availability*

The obvious solution to the problem of lack of availability of suitable data may be in a centralized, accurate, up-to-date and accessible source of property data.

Traditionally government has collected data through a census as well as production and financial data through other mechanisms. As any researcher knows, the data are collected by various agencies and often the same data are collected by different agencies. This lack of coordination and centralisation makes it a nightmare to do research (Carnot *et al*, 2005).

What would happen if departments worked together to share information and established a central database for the micro and macro data required of property researchers? For example, the taxation department may have data on family mortgage payments. The statistics department, using the information data from the taxation department, could produce a median of mortgage payment time series.

Figure 1 illustrates a model for centralized property data collection. A standard measurement and methodology would be required and thus experts and professionals could be involved in compiling property data. In addition, new technologies, such as Internet, would make it possible for enhanced availability and accessibility for researchers and public users.

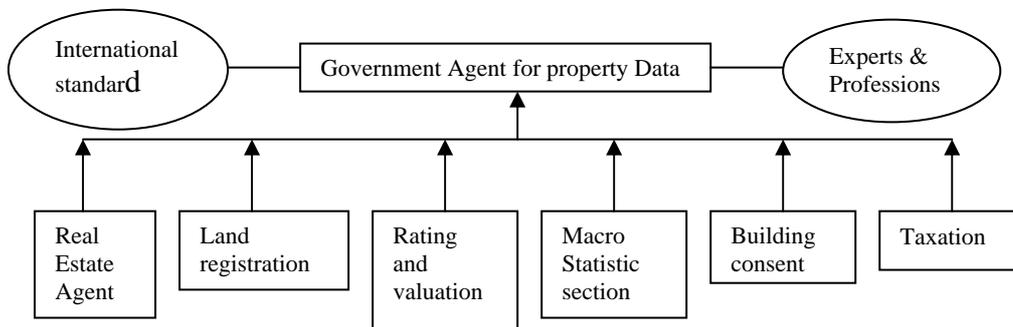


Figure 1: A model of centralized data collection

## Conclusion

This paper has argued that the quality and availability of complete and accurate property data is necessary for all growing economies. The review of the literature indicates that at present raw datasets for analysis are often incomplete and questions of accuracy remain. While solutions to issues of quality are not easily addressed, the issue of data availability seems to be an issue which is more likely to find a solution in the near future.

The impact of growing global economy suggests property research needs to development mechanisms for comparative analysis. As noted above, macro and micro quality datasets are

required to produce models for predicting investment options. The opportunities in property investment suggest that investment performance between cities and countries will be an important factor in the future. The development of a model of a centralised data source for both micro and macro factors would go a long way towards this end.

## Acknowledgement

The research is supported by the Unitec Research and Advanced Practice Committee for 2005 Internal Research & Advanced Practice Fund (URAPC).

## Reference:

- Adair, A, Berry, J and McCreal, S (1995) Property investment in peripheral regions, *Journal of Property Finance*, Vol. 6, No. 2, pp. 43-55.
- Ballard, K (2004) Real Estate Institute of Australia, REIA Update, Issue 1, 9 January.
- Blades, D and Roberts, D (2002) Measuring the non-observed economy, *OECD Statistics Brief*, No. 5.
- Brown, G R and Matysiak, G A (1995) Using commercial property indices for measuring portfolio performance, *Journal of Property Finance*, Vol. 6, No. 3, pp. 27-38.
- Burns, R B (2000) *Introduction to Research Methods*, London: SAGE Publications, p. 336.
- Carnot, N, Koen, V and Tissot, B (2005) *Economic Forecasting*, New York: Palgrave Macmillan.
- Chan, H.L., Lee, S.K. and Woo, K.Y. (2001) Detecting rational bubbles in the residential housing markets in Hong Kong. *Economics Modelling*, Vol. 18, pp. 61-73.
- Cook, L (2004) Revisions to Statistics: their role in measuring economic progress, *Economic Trends*, No. 603.
- Dixon, J & Hamilton, K, (1996) Expanding Measures of Wealth, *Finance and Development*, 33(4): 15-18).
- Dunse, N; Jones, C; Orr, A and Taret, H (1998) The extent and limitations of local commercial property market data, Vol. 16, No. 5, pp. 455-473.
- Fellows, R (2003) *Research methods for construction*, London: Blackwell Science.
- Higgins, D M (2000) The Determinants of Commercial Property Market Performance, PhD Thesis, University of Technology, Sydney, Australia.
- Jenkins, D H (2000) Residential Valuation Theory and Practice, Chandos Publishing (Oxford) Ltd.
- Leishman, C and Watkins, C (2002) Estimating local repeat sales house price indices for British cities, *Journal of Property Investment & Finance*, Vol. 20, No. 1, pp. 36-58.
- Leung, M (2006) Applied research on indices mirrors social and economic trends, *CityU Today*, No. 11, March, pp. 22-27.
- Lum, S K (2004) Property price indices in the Commonwealth – Construction methodologies and problems, *Journal of Property Investment & Finance*, Vol. 22, No. 1, pp. 25-54.
- Mitchell, P M and McNamara, P F (1997) Issues in the development and application of property market forecasting: the investors' perspective, *Journal of Property Finance*, Vol. 8, No. 4, pp. 363-376.
- Morrell, G (1995) Property indices: a coming of age? *Journal of Property Valuation & Investment*, Vol. 13, No. 3, pp. 8-21.
- Newell, G and Chau, KW (1996) Linkages between direct and indirect property performance in Hong Kong. *Journal of Property Finance*, Vol. 7, No. 4, pp. 9-29.
- New Zealand Government website. Available at: <http://www.govt.nz> [accessed 15 April 2006]
- Reichert, A K (1990) The impact of interest rates, income, and employment upon regional housing prices, *Journal of Real Estate and Economics*, Vol. 3, pp. 373-391.
- Tse, R Y C and Love, P E D (2000) Measuring residential property values in Hong Kong, *Property Management*, Vol. 18, No. 5, pp. 366-374.
- Walker A, Chau, K W and Lai, L W C (1995) *Hong Kong in China: real estate in the economy*, Hong Kong: A Brooke Hillier Parker Publication.
- Wyatt, P (1999) The national land and property gazetteer: addressing property properly? *Journal of Property Investment & Finance*, Vol. 17, No. 2, pp. 191-204.
- Wyatt, P (1996) Using a geographical information system for property valuation, *Journal of Property Valuation & Investment*, Vol. 14, No. 1, pp. 67-79.