Obtaining real estate data:
criteria, difficulties and limitations

Stephan V Arthur[1]

Ever since a major section in its 60th Annual Report, the Bank for International Settlements has collected data on residential and commercial property prices. These have been consistently published in "raw" form in its Annual Reports, or been used for various studies in the BIS Economic or Working Papers series, and, more recently, the Quarterly Review. This paper examines the criteria used for collecting such data, the difficulties encountered in compiling them as well as their limitations.

Introduction

In June 1989, the Bank for International Settlements (BIS), in its 59th Annual Report,[2] correlated, in graphical form, the p/e ratio of the Tokyo stock market with the inflation-adjusted price of commercial real estate (land) in six major Japanese cities, noting that Japanese corporations held, at the time, a considerable amount of land. A year later, the BIS devoted a whole section on property markets in its Annual Report,[3] and this was to lead to fairly regular annual publication of the data it had collected in this field. By also correlating both residential and commercial property prices with equity prices in the same section, it was the forerunner for its work on aggregate asset prices.[4] This paper examines the criteria behind collecting and evaluating the data for these real estate data, the difficulties encountered in compiling them as well as the various limitations to the data.

The criteria

There are, generally speaking, for most areas of international data collection, six criteria which need to be satisfied from a statistical viewpoint:[5] regular availability, representativeness, homogenous comparability, unbroken and unchanging description, length of series and data frequency.[6] In a perfect world, data would be always available for all countries under review, have identical measurement parameters, not have any breaks in series, go back to an identical (distant) starting point and, finally, be available on a monthly basis. Alas, the art of statistics (at least in this field) has been to overcome the imperfections found in the real world!

Indeed, the IMF established, in 1996, the Special Data Dissemination Standard (SDDS) to "guide members that have, or that might seek, access to international capital markets in the provision of their economic and financial data to the public. Both the General Data Dissemination System (GDDS)[7] and

---

[1] Statistical Analyst, Departmental Research Assistance, Monetary and Economic Department (e-mail: stephan.arthur@bis.org). Any views expressed are mine and not necessarily those of the BIS.

[2] BIS (1989): Chapter IV, pp 81-2, *59th Annual Report*, June.

[3] BIS (1990): Chapter IV, "Property markets", pp 102-10, *60th Annual Report*, June.

[4] See also "Experience with constructing composite asset price indices" by the author later in this volume.

[5] Such a viewpoint may, incidentally, differ from that of an economist who is involved in a one-off piece of research.

[6] Timeliness is also, usually, an important, if not the most important, criterion for a statistician, but it is not considered here due to the nature of the data. It should nevertheless not be ignored, since even real estate data which is several years out of date is of limited use.

[7] Established in 1997 (author).

the SDDS are expected to enhance the availability of timely and comprehensive statistics and therefore contribute to the pursuit of sound macroeconomic policies; the SDDS is also expected to contribute to the improved functioning of financial markets." The subsequent creation, also by the IMF, of the Data Quality Reference Site (DQRS) is intended to, inter alia, "foster a common understanding of data quality".

Regular availability

First and foremost, data of some sort must be available, and, equally important, be regularly updated. It is therefore only of limited use to the statistician to find a (one-off) piece of research containing data: the best that can be hoped for is that the original data sources are given and that such data can be provided by these sources on a regular basis. Unfortunately, for international institutions such as the BIS, such data collection often has to rely on the goodwill and understanding of a commercial data provider or real estate association, and experience has shown that it is here that the greatest likelihood in the disruption in the flow of data can occur.[8] What sort of data is ideally sought is often tempered by what is found.

By the early 1990s, the BIS had already managed to contact several national statistical institutions and private real estate associations[9] and was able to publish data illustrating the development in residential and commercial real estate for more than a dozen capital cities.[10] Although this was an important starting point, the obvious disadvantages of such a localised focus required further work.[11] In addition, the close correlation between office rents and commercial property prices observed in the 1970s and the early 1980s ended with the office property boom observed in many countries in the late 1980s (and the equally dramatic subsequent decline less than a decade later). The Bank therefore approached other data providers, or went back to its earlier sources, to enquire what else was available.

Representativeness

Data should be representative of present day values, but this is indeed a delicate issue. Just as it is no longer relevant in the industrial countries to include the price of a horse in its consumer price index under the sub-category transport, the price of a house will no longer include an outside toilet,[12] but will probably have several bathrooms, of which at least one may be en suite. Similarly, today's offices will have IT facilities which were in the realms of science fiction a generation ago. Such "upgrades" to indices is common and correct, but what is seen to be representative in one country may well not be the case in another. With 80% home ownership in Spain, for example, any property survey will reveal a very different distribution of types of dwellings to one conducted in a country like Switzerland, where the majority of the population chooses to remain in rented flats. The distribution of the price of these dwellings will also vary accordingly and affect a resultant global index. Equally, amenities will differ

---

[8] Caused, for example, by a change in company policy or by the takeover by another company with a different data dissemination policy.

[9] Inter alia, National Association of Realtors (United States), National Land Bureau (Japan), Ring Deutscher Makler (Germany), Building Societies Association and Department of the Environment (United Kingdom), Associazione Italiana Consulenti Immobiliari (Italy), AN-HYP (Belgium) and Richard Ellis Ltd (for most commercial property prices).

[10] Brussels, Frankfurt, London, Madrid, Milan, New York, Paris, Stockholm, Sydney and Tokyo. For housing prices, Los Angeles and Toronto were also available, and, for office rents, Amsterdam and Lisbon.

[11] For example, commercial property in capital cities are likely to have higher architectural standards, greater functionality and larger, more luxurious meeting room facilities than equivalent property in the provinces. They will therefore command disproportionally higher prices which, in turn, are more volatile than would otherwise be the case in a nationwide coverage. Companies' needs are also subject to greater change, again reflected in commercial property price changes, than is the case for potential homeowners.

[12] Although this was still common in the United Kingdom in the 1950s.

widely: considering the same two countries, the latter will require insulation and heating to keep out the cold in winter, whilst the former will need to provide insulation, and, increasingly, air-conditioning, to keep out the summer heat. Even within a country, what is representative in one region is not at all in another: the majority of housing in the North of England - a traditionally industrial region in the United Kingdom - is terrassed, but detached houses are the standard in the South-West - a part of the country to which many retire.

Comparability

It is unlikely that data for a group of countries will be, in all aspects, comparable. First, as explained above, that which is considered to be representative in one country may not be so in another.

Secondly, the method used to collect data may vary and influence the result. For example, a survey can be conducted by approaching those real estate agents who are members of a national guild: this is unlikely to represent, in some countries, the majority of transactions. Alternatively, the registered lending agencies may be asked for information on housing based on the mortgages they grant: here, cash transactions would escape the net. Notaries could also be approached in those countries where their services are mandatory in finalising property transactions: in order to keep fees down, however, part of the agreed price may be paid in cash and another, lower, price communicated to the notary. Finally, government agencies may provide data, but their original source may be any of the above or be a result of calculations stemming from tax returns.[13]

Thirdly, as hinted earlier, the focus may vary considerably. From a macro-economic viewpoint, and, in particular, for the BIS's recent work on a set of indicators to predict financial crises, a nationwide coverage of property prices is clearly the most desirable. However, especially for commercial property, this is not always available,[14] so that, for several countries, the BIS has to fall back on data relating to prime property in a capital city's centre.[15] Nevertheless, the appearance in recent years of nationwide data[16] indicates that a commercial property price index typically has 80% of the total drawn from property in that country's capital.

Fourthly, even a nationwide index can differ in the way in which it is compiled. A simple average of the prices paid in the individual regions may be taken (Canada) or a weighted average (based on, for example, the population) of the regional survey results (Australia). The index may be the price paid by area rather than unit: this often makes sense for commercial property, but is also occasionally true for residential property (eg France, Italy, Spain).

Finally, there is no guarantee that an index is being compiled at all, but that the series is expressed in, for example, national currency. Indeed, this may well be preferable, since there is a clear loss of information in a simple index.[17] Although this is unavoidable for cross-country comparison,[18] data providers would be advised to keep their series in as "original" a state as possible.

---

[13] For example, in Switzerland, a tax (the "Handänderungssteuer", literally, change-of-hands tax) is levied on each sale of second-hand property, which is based on the sale price, so that an average price is simple to calculate. However, none is levied on new property which, in an under-developed property market (relative to demand) as currently exists in Switzerland, forms a substantial proportion.

[14] Available in the sense that they also satisfy the other criteria.

[15] For example, as provided by Jones Lang LaSalle. These data, however, are also available for a number of Asian-Pacific cities, of which they are, to the best of my knowledge, the sole providers on a collective basis.

[16] Calculated by, for example, Investment Property Databank Ltd.

[17] For example, a table showing the price of a "standard" dwelling (or, better, the price per square metre) for a group of countries in a given year is clearly better than a simple index value which can, at best, only show the relative position - dependent on the base year chosen - or growth rate from that base year.

[18] This raises an interesting point: no attempt has, to my knowledge, been made in comparing exchange-rate-adjusted data. As the BIS has always concentrated on inflation-adjusted property prices, one would assume that exchange rate changes are at least partly taken into account. However, price indices adjusted by either nominal or real (ie inflation-adjusted) effective exchange rates may reveal some interesting differences.

Kennedy and Andersen[19] extended the BIS's coverage of residential property prices to 15 countries but, more importantly, moved from indices of capital cities to nationwide ones.[20] With one exception (Japan: land), the indices referred to real estate (ie the price of building and land). This also reflects the data situation today, but even now, as then, the indices vary in their composition of flats, terrassed or detached houses, single or multiple-occupied dwellings etc.

Continuity

This important criteria, for continuous assessment, requires little explanation, but it is, in reality, one of the most common problems facing statisticians. Apart from the possibility (see "Availability" above) that a source may "dry up" and need to be replaced by another which will almost certainly differ in definition (if not frequency, which is another problem, see below), data from the same source may suddenly change. One reason may be that the source is itself not the primary collector of data, but either the collator or just simply the disseminating body.[21] To be fair, breaks in series often herald an improvement in the data, becoming either more encompassing or moving to a higher frequency (or both). "Splicing" with the previous series, however, remains a difficult problem.

Length of series

This problem is often linked to the previous criteria, since a radical break in series can, if no splicing can be done, considerably shorten a previously lengthy series. Also, when data providers and, by proxy, the Bank, embark on an extended country coverage, it is usually not possible for them to (re)construct a historical series. A table showing the last five years does not, in this respect, pose the same problem as attempts to graph price developments since the 1970s or to carry out historical research analysis (see also below).

Frequency

Precisely in this area of historical analysis, data frequency tends to pose little problem, since most research is done over a fairly long time horizon. Generally speaking, property prices are not thought of as being particularly volatile from one quarter to the next, so that annual data are sufficient. However, a higher frequency is desirable when such indices are used as indicators for monetary or financial stability. In this case, experience has shown that quarterly data best serve the purpose.[22]

---

[19]   Neale Kennedy and Palle Andersen (1994): "Household saving and real house prices: an international perspective", *BIS Working Paper*, no 20, January.

[20]   With the exception of Germany, which was the simple average of four cities (Berlin, Frankfurt, Hamburg and Munich). This has since been replaced by a nationwide index, based on 60 cities.

[21]   This is true, for example, of an increasing amount of the residential property price data used by the BIS. The majority of the countries are now taken directly from its Data Bank, which receives the data from national central banks. They in turn are commonly not the primary sources of this information, but have the data supplied them by the various sources discussed earlier. As a consequence, third-party dissemination can be problematic.

[22]   As can be seen from the Table in the Appendix, the majority of residential property price data are quarterly, whereas most commercial property price data are still annual. Perhaps surprisingly, and indeed problematic given their economic size, residential property prices for Japan, Germany, France and Italy are only available at a lower frequency.

Summing up

As Borio and Lowe (2002) conclude,[23] "The first is *more and better data*. There is, in particular, a remarkable dearth of data on real estate prices, despite their proven role in the genesis of financial crises and, increasingly, in influencing the business cycle. Data gathering has so far been largely left to the initiative of private firms, which naturally tailor the data to their own requirements. Given the 'public good' properties of the data, there seems to be a good case for official authorities to put efforts into this area."

---

[23] Claudio Borio and Philip Lowe (2002): "Asset prices, financial and monetary stability: exploring the nexus", *BIS Working Papers*, no 114, July.

Appendix

The table below show the countries for which the author, in his capacity as statistical analyst, maintains, respectively, residential and commercial property price data, their frequencies, unit (or base period) and start date. Many of the series contain "splices" (see above); where such a link is tenuous (but not necessarily impossible), alternatives are given. Series in square brackets are no longer in use, while those marked with an asterisk are not yet in active production and should be treated with caution.

The Bank's Data Base, referred to earlier, is supplied many other series on residential property prices for nationally-relevant different property types by the reporting central banks; the country coverage does not exceed, however, the list below. In addition, the table only shows data which are felt to best meet the requirements cited in the paper.

Table 1

Real estate prices maintained by the author

Country	Residential property			Commercial property		
	Frequency	Unit/base period	Start date	Frequency	Unit/base period	Start date
Australia	Quarterly	Fiscal 1989 (89Q3-90Q2)	1960 Q1	Quarterly	AUD/m2	1968 Q1
Austria[1]	Semi-annual	1986	1987 H1	-		
Belgium	Quarterly Annual	1953 1953	1981 Q1 1960	Annual (Brussels) Annual (Brussels)[2]	1980 1980	1980 1970
Canada	Monthly[3] Quarterly	CAD 1980 Q4	1980 M1 1970 Q1	Quarterly (Toronto)	1985	1985 Q1
China	Annual*	CNR/m2	1987	Annual*	CNR/m2	1987
Denmark	Quarterly	1980	1970 Q1	Annual (Copenhagen) [Semi-annual][2]	1984 Q3 1980 H2	1982 1965 H1
Euro area	Annual	2000	1991	-		
Finland	Quarterly Annual	1983 1970	1978 Q1 1970	Annual Annual (Helsinki)	pcpa EUR/m2	1998 1971
France	Semi-annual Annual	EUR/m2 1997	1995 H1 1960	Annual Annual (Paris) Annual (Paris)[2]	EUR/m2 1980 1980	1986 1980 1970
Germany	Annual Annual[2]	2000 DEM/m2	1975 1971	Annual Annual (Frankfurt) Annual (Frankfurt)[2]	pcpa 1980 1980	1996 1980 1971
Greece	Quarterly*	1997	1997 Q1	-		
Hong Kong	Monthly Quarterly	1999 1999	1993 M1 1980 Q1	Quarterly	1999	1988 Q1
Ireland	Quarterly Annual	IEP	1976 Q1 1970	Quarterly Annual	1982 1982	1994 Q4 1982
Italy	Semi-annual	EUR/m2	1988 H1	Annual (Milan)	1983	1983
Japan	Semi-annual	1990 M3	1955 H1	Semi-annual	1990 M3	1955 H1
Korea	Monthly*	1995	1986 M1	-		
Malaysia	Annual*	1990	1988	-		
Netherlands	Monthly Annual	EUR 1980	1976 M1 1965	Annual Annual (Amsterdam) Annual (Amsterdam)[2]	pcpa 1980 1980	1995 1980 1970

Table 1 (cont)

Real estate prices maintained by the author

Country	Residential property			Commercial property		
	Frequency	**Unit/base period**	**Start date**	**Frequency**	**Unit/base period**	**Start date**
New Zealand	Quarterly Quarterly	1999 Q3 Fiscal 1971 (71Q3-72Q2)	1989 Q4 1962 Q2	Semi-annual	NZD/m2	1980 H1
Norway	Quarterly Annual Annual	2000 2000 1969 (?)	1991 Q1 1980 1970	Annual Annual (Oslo) Annual (Oslo)[2]	pcpa 1990 1980	2000 1990 1970
Portugal	Monthly	1988 M1	1988 M1	Annual*	pcpa	2000
Singapore	Quarterly	1998 Q4	1988 Q2	Quarterly	1998 Q4	1988 Q2
South Africa	Monthly[3]	2000	1980 M1	Annual	pcpa	1995
Spain	Quarterly Annual (Madrid)	EUR/m2 ESP	1987 Q1 1975	Annual* Annual (Madrid)	pcpa 1980	2001 1980
Sweden	Quarterly	1980	1970 Q1	Annual Annual (Stockholm) Annual (Stockholm)[2]	pcpa 1980 1980	1984 1980 1970
Switzerland	Quarterly	1970 Q1	1970 Q1	Quarterly	1970 Q1	1970 Q1
United Kingdom	Quarterly Monthly	2001 Q1 1983	1968 Q2 1983 M1	Annual [Monthly]	1980 1986 M12	1970 1986 M12
United States	Quarterly [Monthly]	1980 USD	1975 Q1 1968 M1	Quarterly	1977 Q4	1977 Q4

[1] Vienna only; discontinued at end-2002.   [2] Confidential proprietary data.   [3] Seasonally adjusted.