

Data Sources

Introduction

9.1 In practice, because of the high cost of undertaking purpose-designed surveys of house prices, the methods adopted by statistical agencies and others to construct residential property price indices have mainly made use of administrative data, the latter usually being a function of the house price data sets generated by a country's legal and administrative processes associated with buying a house. The indices so constructed can vary according to the point in the house purchasing process at which the price is measured. For example, the final transaction price or the earlier valuation used for securing a loan could be used as the "price" of the property. Furthermore, different administrative data sets will generally collect information on different sets of characteristics associated with the sales of the properties. These differing information sets will generally affect index compilation methods, often acting as a constraint on the techniques available to quality adjust for houses of different sizes, locations, etc. Thus data sets have historically acted as a constraint on index construction.

9.2 This chapter examines the different sources of data used for constructing residential property prices indices. Although it focuses mainly on price data, the chapter also considers how the choice of weighting scheme can be constrained by the information generated from the house-purchasing process. Different weighting schemes, notably whether an index is stock or sales weighted, produce price indices which measure different concepts. In these circumstances it is important that there is a clear understanding of what the target measure is so that the indices compiled can be evaluated against the target measure to determine fitness-for-purpose.

Prices

The Process of Buying and Selling a House

9.3 The process of buying and selling a property normally takes place over a period of several months or more. The particular stage in this process at which the price is entered into an index will depend on the source of the data and this has consequences for what is being measured and for the comparability of different indices. Price data for a residential property price index may be taken at the following stages:

- As soon as the property is on the market (advertised or asking price). Typical data sources: newspapers, real estate agents.

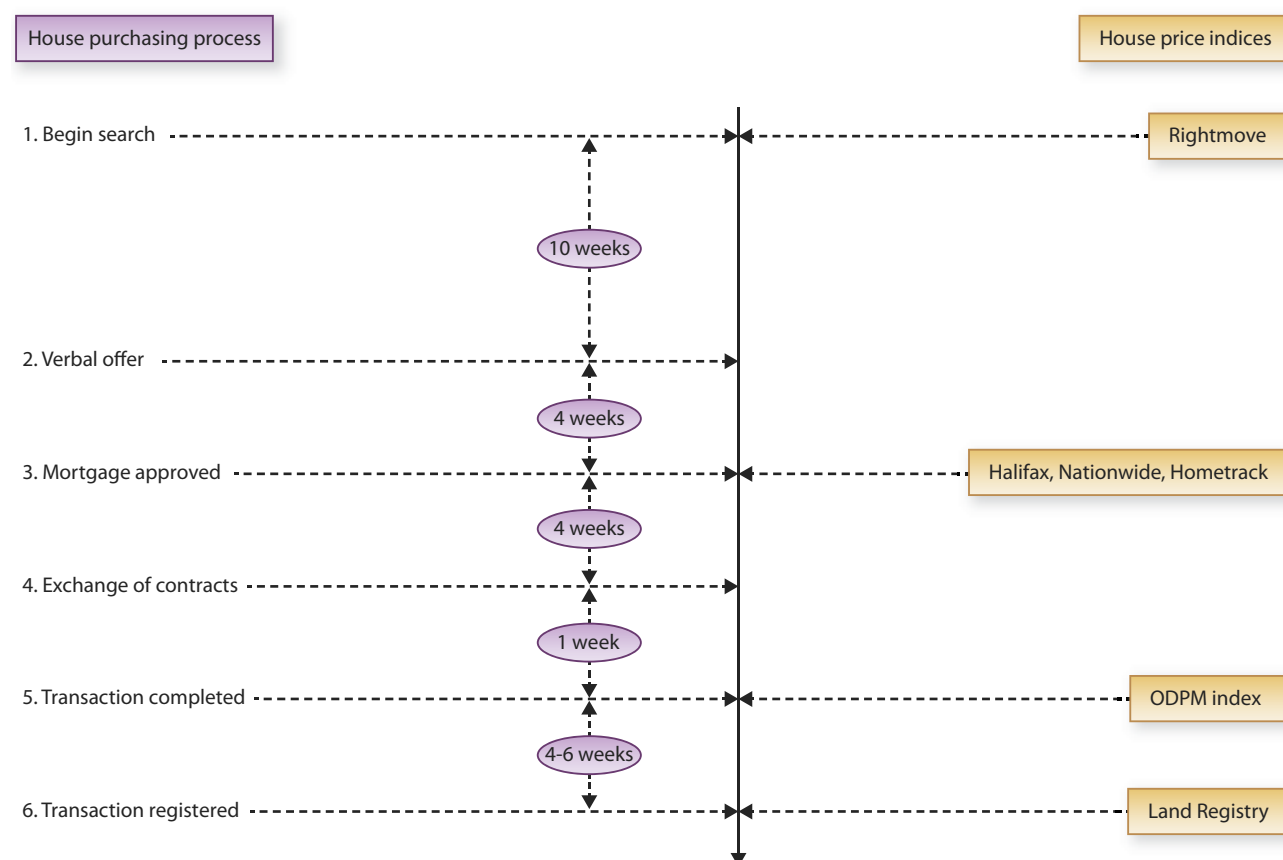
- Mortgage applications. Typical data source: mortgage lenders.
- Mortgage approved. Typical data source: mortgage lenders.
- Signing of binding contract. Typical data source: lawyers, notaries.
- Transaction completed. Typical data sources: land registries, tax authorities.

9.4 Each source of price data has its advantages and disadvantages. For example, a disadvantage of advertised prices and prices on mortgage applications and approvals is that not all of the advertised prices will end in transactions, and the price may differ from the final negotiated transaction price. These prices are likely to be available sometime before the final transaction price. Indices that measure the price earlier in the purchase process are able to detect price changes first, but they will measure final prices with error because prices can be renegotiated extensively before the deal is finalized.

9.5 It should be noted that the availability of different sources of price information at different points in the buying and selling process can be an advantage. For instance, changes in the relationship between asking price and selling price may provide an early indication of a change in the housing market. The diagram below illustrates the situation in the UK; see also the case study for the UK in Chapter 10.

9.6 Most data sources are susceptible to all the disadvantages of using administrative systems for statistics. The use of administrative data in economic statistics has been associated with four challenges: definitions, coverage, quality, and timeliness – with expected trade-offs against compilation costs. Definitions and coverage are sometimes placed under the one heading of "coverage": to embrace the types of units covered and the degree of coverage. For example, cash sales could be recorded but properties bought with a mortgage may not be covered or some cash sales may not be recorded if, for example, they are under the threshold for tax liability.

9.7 The underlying problem arises from the fact that the data are primarily recorded as a step in the administrative process and not as an input into a statistical system. The data are not under the control of the statistician. The inherent weaknesses in administrative data need to be taken into account when using the data and in interpreting the results, in particular when they are used as a substitute for statistical data rather than as a supplement to or in conjunction with purpose-designed statistics. Some of the weaknesses may be overcome by an appropriate methodology, such as combining complementary data sources, and possibly by using some form of modeling.

Diagram: House purchase timeline and house price indices

Source: Bank of England and former Office of the Deputy Prime Minister (ODPM)

9.8 A number of basic characteristics come into play in considering the suitability of different data sources.

- **Definition.** This is closely associated with conceptual issues and what the target measure of an index is.
- **Coverage.** Issues relating to coverage will be determined by the operational boundaries of the agency or business providing the housing data. For example, the agency could cover country-wide property sales or just cover a particular region or the transactions covered could relate only to cash purchases or to properties purchased using a mortgage loan. For a government agency, the operational boundaries will be dictated by the regulations and legal processes involved with the purchase of residential property. Inevitably, for public and private data providers coverage will also be heavily dependent on the resources at the disposal of the agency or business and its efficiency in providing data. All these factors are outside the control of the index compiler and can impact on data quality and on any divergence between intended coverage of the residential property price index and actual coverage.
- **Quality.** When considering the issue of data quality, it should be borne in mind that the administrative authority is likely to focus on validating the information which is pertinent to the sale and to the execution of its duties and which reflect the laws and regulations which it is required to comply with. There may be other information which is collected which is of interest to the statistical agency, but which is only of limited relevance to the administrative authority. For instance, this may be the case for some house characteristics which the statistical agency may wish to use for quality adjustment. At the end of the day, the reliability of administrative data will depend on the incentive for data suppliers to give correct information and complete information. There can be mutual advantage to both parties from the statistical agency helping the administrative authority to improve the quality of its data. This can be done by giving feedback on the consistency of data entries and from advising on more general weaknesses. Some statistical agencies provide the administrative agency with an incentive to improve their data collection by compiling custom-designed statistics for the data supplier in return for access to the raw data.
- **Timeliness.** The timeliness of administrative data will depend on who is responsible for reporting to the

administrative authority and on the incentive for timely reporting. For instance, there may be a big incentive for a buyer to obtain approval from the mortgage company, for a house loan and for the mortgage company, to quickly get an accurate and up-to-date valuation so that the sale can go through, with all parties safeguarded, before another potential purchaser takes an interest in the property. On the other hand, there may be less of an incentive to register the sale quickly with the official land registry once completed.

One of the keys to the successful use of administrative data is to have an intimate and detailed knowledge of the data collection processes and associated operational systems.

9.9 Each source of price data is considered separately below. Where more than one data source is available to the index compiler, the opportunity arises for consistency checks and for data from different sources to be combined. For instance, it may be possible to use the property valuations carried out for the approval of loans to predict the final transaction price recorded much later on by the land registry. This depends of course on the stability of any correlation found between the two.

Seller's Asking Price: Estate Agents, Newspapers, Etcetera

9.10 Information on the seller's asking price can be collected through surveys of real estate agents or from an examination of advertisements in newspapers, magazines or online. One of the main advantages of indices constructed from such information is their timeliness. By taking asking prices, indices constructed using this information can provide a timelier estimate of house prices than those indices that are based on subsequent transactions. They also have an advantage over house price indices based on information from mortgage lenders, as the latter are limited to transactions involving mortgages. However, indices based on initial asking prices have a major drawback. Houses can be withdrawn from market and the agreed selling price may not equal the seller's asking price. These indices ignore reductions in prices that sellers subsequently make, for example when the housing market is on a downturn, or offer prices above the asking price when the housing market is buoyant. Such indices can therefore present an over-optimistic outlook when the housing market becomes depressed and an over-pessimistic outlook when the housing market is recovering. The fact that they cannot be relied upon to present an accurate picture of the housing market in the short term devalues their usefulness to most users, most particularly those interested in the early detection of turning points in the housing market or an advanced indicator of the future direction of house prices. It should be noted that the differences between initial asking price compared

to actual transaction price also imply that the calculation of "average house price estimates" can sometimes be misleading.

9.11 Information collected on a seller's asking price cannot always be easily verified and, as well as depending on a balanced and representative sample, relies on the honesty and knowledge of those being surveyed and when drawn from advertisements, the accuracy of the information, especially when it is from a website. For example, it has been argued that real estate agents are more likely to be optimistic about prices and have a vested interest in prices going up rather than down and that this may influence survey results. On the other hand, an estate agent might suggest to a seller an unrealistically low asking price in order to get the property off their books quickly to get the commission. It has also been argued that websites will tend to be biased towards properties that have a competitive asking price to entice potential sellers. All this is, of course, speculation but it does bring home some of the potential difficulties associated with these sources.

9.12 Surveys of real estate agents have some inherent advantages over surveys of advertisements. Agency surveys can be based on a more scientifically selected sample and can provide information on a representative selection of those properties on the market, including those which typically are not covered in advertisements. Data from real estate agents might include extensive information on the characteristics of the property and this information is extremely important for quality adjustment (using either hedonic regression methods or stratification methods as was seen in previous chapters). Also the survey questionnaire could collect information on issues such as: what is the average selling time or what has been the recent difference between asking prices and selling prices (e.g. "higher" or "lower") or on the number of potential buyers registering and the number of properties listed with the agent. This information can help put the price information used in compiling the index into context and can be useful for interpretation of the final results. But such surveys typically do not record the asking price of a specific property. Rather, the questionnaire would normally ask the real estate agent to give the "average asking price" for a selection of representative properties.⁽¹⁾ For example, this might be for each of four standard property types (flat, terraced, semi-detached and detached) in a number of different locations. It is this information which is used to create an average property price for each property type in each location, which is used in turn to compile the corresponding price index. In contrast, the inherent advantage of a survey of advertisements is that the latter will collect the actual asking price for each of the advertised properties.

⁽¹⁾ Some surveys also ask for "achievable" price and use this to construct a house price index.

9.13 In summary, although a house price index based on surveys of asking prices may be more timely, the difficulties in determining exactly how the survey information was compiled and the uncertain relationship between asking price and selling price mean that care should be taken if such an index is to be used as a barometer of house prices.

The Initial Offer Price Accepted by Seller: Mortgage Companies

9.14 Many countries turn to mortgage lenders as the main data source for their house price index. The information is stored in the lender's computer system and serves the operational business needs of the mortgage lenders. This database may include the initial offer price made by the potential purchaser, the valuation price used for authorising a loan and sometimes also the final transaction price. Information from mortgage companies can suffer from all the disadvantages of using data drawn from administrative systems, as described above, but these databases can be a rich source of timely information.

9.15 However, data from mortgage lenders suffer from a major drawback: they exclude non-financed home purchases. Research has indicated that cash buyers account for about a third of the UK market and cash buyers tend to purchase either very cheap or very expensive properties. This would not be problematic if it was not for the fact that dwellings purchased for cash can experience different price developments compared to those financed by a mortgage. This is likely to be particularly the case at turning points in the market where different ends of the housing market may react differently to the economic circumstances and the premium for a cash-buyer increase. For instance in a down-turn, people at the top end of the market who were considering selling their homes to release equity may hold back from putting their homes on the market at a reduced price, so the supply of houses for sale falls and is mainly from owners who, for one reason or another, are very keen to sell. However, at the same time the number of active potential mortgage-based buyers could drop significantly as people are reluctant to take out larger mortgages. But some people will need to sell. In this situation a cash-buyer for a house at the upper end of the market will be in a relatively stronger position to negotiate a bargain price than in a more stable market.

The Valuation Price for a Loan: Mortgage Companies

9.16 Mortgage companies will obtain an independent valuation of a property before approving a loan. The valuation that the mortgage company provides the customer with at the time of the mortgage approval can be some weeks after the buyer and seller have negotiated a final

price and the buyer has made the initial application for a loan. In practice there is a negotiation process between these two stages in which it is possible for the agreed purchase price of the dwelling to change. This can be the case when the independent valuation differs from the price the purchaser and buyer had agreed upon or where the purchaser has paid for a detailed survey of the property which reveals that substantial repairs are necessary. For instance, it is fairly common for a buyer to try to leverage a price reduction if the valuation by the mortgage company turns out to be significantly lower than the previously agreed price, or if a survey of the condition of the property reveals the need for new roofing. Clearly, the difference between the initial offer price and the follow-up valuation and any process of re-negotiation which takes place subsequently can result in the measured rate of house price inflation to differ from the true rate as measured by the actual transaction price.

9.17 The house price change measured by indices based on valuations by mortgage companies⁽²⁾ can differ from the price change shown by the offer price and both may differ from the price change based on final transaction prices even when taken from the same sample of mortgage lenders. Thus, it is important to understand exactly what an index is measuring.

The Final Transaction Price: Mortgage Companies

9.18 The time lag between the mortgage application, mortgage approval and purchase completion stages and the differences in the corresponding values of the house prices illustrate the trade-off between timeliness and accuracy. The final transaction price is not always recorded by mortgage lenders and is often extracted instead from legal records such as entries made in land registers, which additionally also include sales that did not require a mortgage. But there can be a long time lag between the completion of the transaction and the recording of the sale in the land register. One of the main advantages of data from mortgage lenders is its timeliness. Initial offer prices and valuations provide an earlier indication of current prices, as these data are available earlier, and final transaction prices may be available sooner from the mortgage lender than from the land registry. It is for this reason that the exploitation of information from mortgage lenders on final transaction price may be a preferred option. The final transaction price held by mortgage lenders can be easily verified against land registry records to alleviate any concerns regarding accuracy and credibility.

⁽²⁾ It has to be taken into account that prices from mortgage valuations, like prices based on any valuation, depend on the objectivity of the evaluation process. Thus, it has been mentioned that the mortgage valuations can sometimes be influenced by the credit policy of the bank, indicating potential difficulties associated with these sources.

The Final Transaction Price: Administrative Data from Property Registers and Tax Offices

9.19 Ideally a house price index would be based on actual transaction prices at the time when the property is sold and the sale completed. The signing of the first binding contract best fits this requirement because of its timeliness but in practice there can be some ambiguity about the point at which a contract is binding, e.g. whether this is at the point where an offer is formally accepted (e.g. when sealed bids are opened), or when a contract is signed or when the contract is exchanged. Similarly, there can be a difference between when a contract is signed and when the transfer of ownership takes place and when it is recorded in the property registers or at the tax office.

9.20 In theory, information from property registers or tax offices will cover all properties, including cash purchases as well as purchases via a mortgage and thus these databases should be the most comprehensive of all the sources available to the index compiler. But, in practice, comprehensiveness cannot be guaranteed, particularly if there is a disincentive for the owner to register a property. For example, when the primary purpose of registration is for taxation purposes, properties may not get registered at all, or may be registered with some relevant detail such as square metres of floor space missing or incorrectly recorded, in order to avoid tax or reduce the tax charges.⁽³⁾

Valuation Price for Taxation and Payment for Local Services: Tax Offices

9.21 In many countries, the central or a local government may impose a monthly or annual tax or service charge on residential properties, for funding the provision of public services such as road maintenance, police and fire services or refuse collection. In many cases, the tax bill faced by an individual is proportional to the assessed value of property and the latter is usually based on a valuation undertaken by professional chartered surveyors either under contract or directly employed by the taxation authority. The valuations should take into account characteristics of the property, such as location and size of plot. However, they rely on accurate information about the properties and also on the chartered surveyors' assessments, which are difficult to verify. Also the updating of the valuations tends to be infrequent due to the field costs involved. Because of these drawbacks, the information collected can sometimes be of limited use in the construction of residential property price indices.

⁽³⁾ There is a related problem: the transaction price may not be a market price because the transaction, while genuine, is between relatives or friends. For example, parents may decide to pass on the family home to their children at a below market price.

That said, this source of official valuation information has been exploited by statistical agencies; see the material on the SPAR method of index construction described in Chapter 7.

Other Expert Opinion Information: Surveys of Estate Agents Organisations, other Professional Bodies and their Members

9.22 In some countries, regular surveys are conducted of real estate agents, chartered surveyors or their corresponding professional bodies, asking about house prices and housing stock. These "opinion" surveys are typically restricted to asking respondents to give a view on whether house prices are moving up, down or flat. These surveys do not give an indication of how much houses are worth or by how much prices are falling or rising but they can provide an up-to-date and broad-based picture on the direction of price change in the housing market to supplement and help to add credibility to the latest figures from a residential property price index. For instance, a significant change in the difference in the proportions of real estate agents who think prices are going up and those who think prices are going down might provide an early indication of a change in the housing market not yet detected by the currently available statistics on mortgage lender valuations. Contextual information of this kind adds value and is regularly used by commentators when interpreting official house price indices.

Evaluation of Data Sources for Fitness-for-Purpose

9.23 The overall usefulness of the above sources of information on residential property prices will very much depend on their fitness-for-purpose for the particular applications to which they are being used. To gauge fitness-for-purpose requires an evaluation of the intrinsic advantages and disadvantages of the index against an agreed set of criteria, i.e. an evaluation against user needs.

9.24 Chapter 2 reviewed the many different uses of house price indices: as a macro-economic indicator of inflation; for monetary policy targeting; as a measurement of change in wealth; as a financial stability indicator to measure risk exposure; as a deflator for the national accounts; as an input into an individual citizen's decision making on whether to invest in residential property; as an input into other price indices, in particular the Consumer Price Index (CPI), and for use in wage bargaining or indexation.

9.25 An effective evaluation of the different sources of data on house prices is dependent on a systematic analysis of user requirements. User needs have a significant impact on decisions relating to the conceptual basis of an index and the associated statistical requirement. This may take the form of a series of questions reflecting the different reasons why users may want information on house prices. For instance, whether an index of house prices is to be used as one of a suite of general macroeconomic indicators, as an input into the measurement of consumer price inflation, as an element in the calculation of household wealth or as a direct input into an analysis of lenders' exposure. Such an analysis can then be transformed into a statistical user requirement and an associated conceptual framework by expressing the needs in statistical terms and identifying the common linkages and corresponding relationships at a micro and macro level. The different data sources can then be evaluated against the statistical need.

9.26 The following list of desirable properties for a residential property price index constitute a possible set of criteria for an evaluation of alternative data sources for fitness-for-purpose for different uses.⁽⁴⁾ The list builds upon the discussion at the beginning of this chapter. The relative importance of each of the criteria will depend on use and in essence constitutes a statistical requirement. There will also be the usual trade-offs between fully meeting user needs and the costs of data collection.

Definitions and Measurement Concept

9.27 This also covers *coherence* with other statistical outputs. It represents the user requirement at the most basic level. Consider the needs of governments and analysts looking at inflationary pressures and those with a direct investment in real estate. The primary focus of these users may be the cyclical nature of prices and the ability of real estate prices to lead to destabilising booms and slumps in the economy as a whole. For this purpose, users will be looking to a variety of indicators, including indices of the volume and price of real estate transactions, as well as macro-economic indicators for modelling the economic cycle and predicting peaks and troughs. Analysts looking at the inflationary pressures of real estate price rises in comparison to other price rises may be interested in including in a CPI the inflationary costs of owner-occupier housing costs by means of a house price index based on the net acquisition cost basis but excluding land.

9.28 For users wanting a general macro-economic indicator, an index based on all purchases – both cash and those with a mortgage – is appropriate. Taking transaction prices

solely from data supplied by mortgage lenders represents a serious deficiency. Conceptually, land registry data would represent a better source as it should cover all transactions. The challenge is to find a source of price data which readily fits, or can be manipulated to meet, the requirements of users interested in the inclusion of owner-occupier housing costs in a CPI on a net acquisition cost basis, that is, excluding the price of land.⁽⁵⁾

9.29 In contrast, users interested in an analysis of the current value of the real estate portfolio against which outstanding mortgages are secured, will require an index of changes in the price of the properties for which mortgages were issued, weighted by the amounts loaned for each type of property at the time at which they were issued. For both of these measures, the value of the land underlying the buildings is as important as the value of the buildings themselves and it is the total value of the land and buildings which is of interest. For these users, data from mortgage providers on property prices and the size of new mortgages and outstanding debt will fit the purpose.

9.30 Now consider the needs of employers and trade unions when negotiating wage settlements. Their primary focus will be the effects of price changes on the standard of living of workers. For this purpose users will be looking to a CPI that includes the cost of keeping a roof over their heads – for owner-occupiers the cost of mortgage interest payments and the repairs costs. The measurement of this will require the calculation of the mortgage outlay at time of purchase and the subsequent repayment history will need a sales weighted house price index. In an ideal world re-financing would be excluded. The repairs element may be measured by the calculation of depreciation. For this, a stock-weighted smoothed house price index is most appropriate. In addition, there is the issue of land where it is often argued that in most circumstances land is an investment which appreciates and that its inclusion in a depreciation calculation is inappropriate.⁽⁶⁾ Thus an index excluding the price of land may be required.

9.31 For the calculation of mortgage outlay, the user can again rely on information supplied by mortgage lenders, but not for the estimation of depreciation, where the value of land may again need to be separately identified.

9.32 As a final example, consider the needs of national accountants, who are seeking appropriate deflators for national accounts. Their needs again will be different. Real estate appears in the National Accounts in several ways (for details, see Chapter 3):

⁽⁴⁾ See also Chapter 3 where a listing of user needs is presented based on discussions between users of house price indices and the Office for National Statistics. In that section, it was pointed out that there is a trade-off between the desires of users to have a family of more detailed indices (stratified by location and type of housing) and the quality of the indices: more detail inevitably leads to less accurate indices.

⁽⁵⁾ In most countries for most transactions, land and building are purchased together as a "single package", so the two components are typically not separated in the information generated by records relating to the transfer of ownership. As such separating the prices would require a supplementary exercise. In Chapter 8 it was outlined how hedonic regression can be used to decompose the overall price index into land and structures components.

⁽⁶⁾ There are other more general issues, which are not addressed here, to do with the measurement of depreciation and its inclusion in a consumer price index.

- The imputed rental value received by owner occupiers for buildings is part of household final consumption.
- The capital formation in buildings, as opposed to land, is part of gross fixed capital formation, depreciation, and the measurement of the stock of fixed capital.
- Land values, which are an important part element of the national stock of wealth.

In each case the derivation of volumes from values requires price indices for respectively: the imputed rent of owner occupied dwelling units weighted by the stock of different types of owner occupied housing; new house purchases weighted by the transactions in new houses but excluding the land component; and of the whole housing stock including land weighted by the housing stock

9.33 It can be seen that user needs will vary and that in some instances more than one measure of house price or real estate inflation may be required. It can also be seen that coherence between different measure and with other economic statistics is important and that achieving this will be especially difficult as statisticians are unlikely to have an ideal set of price indicators available to them.

Coverage

9.34 Coverage includes not just whether all properties are covered irrespective of whether the property is owned outright or being funded by a mortgage but also whether country-wide property sales or valuations are covered or just those in a particular region and whether all price ranges are covered. It can be noted that even where the primary need is for a national index, regional indices can be in demand for analytical purposes. House price information from any individual mortgage lender is unlikely to be representative of the country as a whole, not only because of the exclusion of cash purchases but also because lenders often focus their business on particular regions.

Quality

9.35 Quality relates to the accuracy and completeness of the information, i.e. there are no serious errors and the information is what it purports to be. Compared with other administrative data, house price information from a land registry is likely to score relatively highly in terms of accuracy due to the legal requirements to record property transactions and exchanges of ownership. However, the reliability of data from any administrative source is difficult to validate.

Timeliness

9.36 Indices that measure prices earlier in the purchasing process are able sooner to detect price changes and turning points in house price inflation. This is likely to be

particularly important when used, say, for macro-economic policy and monetary targeting but less important for a national accounts deflator. Data from mortgage lenders may better suit the needs of those engaged in macro-economic policy and monetary targeting, even though cash purchases are excluded, whilst land registry data may better suit the needs of, for example, those calculating deflators.

Detail for Quality Adjustment and Mix-Adjustment

9.37 This relates to two (related) issues: the degree to which residential property price indices are able to adjust for changes in the mix of properties sold and to eliminate the effect of quality changes of the individual dwellings. For this purpose, “real time” information is needed on price determining attributes such as size of plot, size of house, type of property (flat, house, semi-detached or detached), location, the condition of the property, whether it has central heating, a fully-fitted kitchen and bathroom, etc. Quality (or mix) adjustment is essential in order to construct an accurate price index for housing components.⁽⁷⁾ It is unlikely that any of the sources of prices data listed above will be ideal for all purposes. The amount of detailed and relevant characteristics data will depend on the individual data set.⁽⁸⁾

Frequency

9.38 Frequency essentially relates to how frequently an index can be computed, e.g. once a month or once a quarter. There is a tradeoff between frequency and accuracy. For a particular geographic area and type of housing, current information on the price of houses in a given strata will come from sales of old and new houses in that strata during the chosen time period. If the frequency is chosen to be a month as opposed to a quarter, the monthly sample size will only be approximately one third of the quarterly sample size. Thus a monthly house price index based on sales of properties in the given strata will be subject to increased sample volatility (and hence will not be as accurate) as compared to the corresponding quarterly index. Volatility of a monthly index *may* be reduced by making the strata “bigger”,⁽⁹⁾ e.g., different neighbourhoods could be combined within the same general location but this leads to another tradeoff between fineness

⁽⁷⁾ The various methods available for constructing quality adjusted house price indices were discussed in Chapters 4-8.

⁽⁸⁾ In cases where the real estate agent data base includes the final selling price of the listed properties along with the main characteristics of the properties, this information base is probably the “best” for most purposes. However, the sample of listed properties needs to be compared with the properties listed in land registry offices to ensure that the coverage of listed properties is adequate for the purpose at hand. When constructing price indices for the stock of housing, it will be necessary to have census information on housing stocks along with post census information on demolitions and the construction of new dwelling units.

⁽⁹⁾ It is not certain that combining strata will reduce index volatility if house prices in the different micro strata have different trends.

of the strata (which many users may want) and accuracy of the index (which all users want).

9.39 It may be possible to provide smoothed monthly house price indices that are say a three month moving average of the raw monthly indices⁽¹⁰⁾ or the statistical agency could provide both monthly and quarterly indices and let users choose their preferred index.⁽¹¹⁾ It is not possible to provide definitive advice on how frequent a house price index covering a certain stratum should be published. The issue of frequency must be decided by the national statistical agency, taking into account user needs and data availability.

Revisions

9.40 Revisions can refer to either revisions resulting from subsequent returns (so that the series itself is revised) or from other sources of more relevant data subsequently coming on stream (so an early indicative measure is eventually replaced by a precise measure of what needs to be measured).⁽¹²⁾ For instance, an example of the former might be revisions arising from late registration of property sales. An example of the latter might be where an initial offer price recorded on the mortgage application form is used as an early indication of movements in transaction prices but is subsequently discarded when land registry data on actual transaction prices (which takes into account any price renegotiation before the sale is finalised) eventually comes on stream at a much later point.

9.41 The extent to which figures are revised due to the receipt of subsequent returns is partly determined by the reference point of the prices data and partly by the point in time when the particular data set is received by the statistical agency: the earlier is the data reference period in the purchasing cycle and the earlier the particular data set is received, the more the index will be subject to revision. Thus, although information from the registration of property sales is appropriately referenced and provides a definitive source of information on property prices, the time delay that can sometimes take place in some countries for the legal registration of property transfers can mean that the register is not final until, say, twelve months the sale of the property.

⁽¹⁰⁾ The Australian Bureau of Statistics makes frequent use of this technique for a wide range of its statistics. If the window length is 12 months, then the resulting smoothed index can be regarded as a seasonally adjusted index, centered in the middle of the 12 month period under consideration. For a variant of this smoothing technique, see Chapter 4.

⁽¹¹⁾ There is a possibility that some users may be confused by having more than one index covering essentially the same housing strata. However, the Bureau of Labor Statistics now has two monthly published Consumer Price Indices: their headline Lowe type CPI which is not revised and a second index which is an approximation to a superlative Törnqvist index (which is revised). Users in the U.S. seem to have accepted multiple indices in this context.

⁽¹²⁾ A related issue is that some of the methods for constructing an RPPI, such as the multiperiod time dummy hedonic method (see Chapter 5) and the repeat sales method (Chapter 6) suffer from revision in the sense that previously computed figure will change when new data is added to the sample. In some cases, revised indices are published while in other cases, the rolling window technique with updating due to Shimizu, Nishimura and Watanabe (2010) and Shimizu, Takatsuji, Ono and Nishimura (2010) is used. The rolling window with updating technique does not revise the historical index up to the current period.

9.42 Valuation prices kept by tax offices for taxation and payment for local services and the final transaction price recorded by mortgage companies are least likely to be subject to revision, whilst the final transaction price based on administrative data held on property registers and tax offices could be subject to revision over a long period depending on the time-lags involved in the legal processes of recording changes in ownership.

Comparability

9.43 Comparability refers to the degree of *inter-country comparability* between house price indices. This is important because comparing house prices from non-harmonised national data can be problematic as differences in concept, index construction, market coverage, quality adjustment procedures, etc. can make cross country comparisons difficult. Differences in frequency, timeliness and revisions policy can also cause comparability problems.

9.44 Problems can arise at both the national and international levels:

- Users in individual countries can be confronted either with a lack of relevant statistics or with different statistics for different time periods and with varying time-lags and these statistics can be based on different data sources or compilation methods.
- For users seeking international comparisons the situation is complicated by significant differences among countries with regards to the availability of data and the challenge this represents for compiling like-for-like comparisons and interpreting relative trends among countries. The complication of aggregate price indices covering groups of countries – a requirement for co-ordinated economic policy and monitoring across an economic area such as the Eurozone⁽¹³⁾ – is a further challenge.

From Chapter 10 it can be seen that the methods employed for the compilation of residential property price indicators vary considerably *between* countries, and even between alternative sources *within* individual countries.

Weights

9.45 The data sources drawn on for the weights in a residential property price index are a function both of the data needs of the target index and of the availability of the required information. Also the data needs depend not only on the conceptual basis of the index but also on detailed aspects of index construction, such as the method of quality adjustment and any subindices that are required

⁽¹³⁾ Consisting of the seventeen member states of the European Union that have adopted the Euro as of 2012.

for analytical and other purposes. For instance, the construction of a mix adjusted property price index based on transactions requires that enough information is known about the sales in each period for them to be classified into groups sufficiently homogenous so that the unit values can be treated as prices. In the housing market, the problems are compounded by the low volumes of sales for certain house types in particular geographical areas which could lead to many cells being empty.⁽¹⁴⁾

9.46 Putting these detailed issues of construction to one side, the conceptual basis of the index is the main factor determining the data needs relating to weights. One price index cannot meet the diverse needs of users. For estimating gross capital formation, for instance, only new houses should be included while estimating the effect of price changes on capital stocks requires the index to cover all transactions.

9.47 The weights can be derived from a number of sources, in particular, from national accounts data, periodic national censuses which collect information on the housing stock, information from banks on the loans taken out for house purchase, construction statistics, official registers recording ownership, etc. There can be a lack of coherence between these different data sources resulting from the long and quite often involved processes associated with buying and selling a house and the fact that a valuation or offer price associated with an application for a mortgage will not necessarily lead to a sale and change of ownership. Other issues arise also, such as the distinction between what is being built for selling and what is being built for renting out. This sort of information is rarely readily available from one statistical source. It is for this reason that the construction of weights may draw on a multitude of different sources.

Developing Countries, Traditional Dwellings and the Informal Housing Market

9.48 For many developing countries, a significant proportion of the housing stock consists of newly constructed buildings on family owned land or of old buildings which have been significantly upgraded since they were first constructed. There can also be a significant element of owner-constructed housing. Construction may take many years and at any point in time a substantial proportion of the

houses could be considered incomplete. The use of formal mortgage finance is often very limited but informal finance may be used. House construction can vary from shanties built on compacted soil with salvaged materials to substantial multi-room dwellings built on concrete foundations with concrete blocks. Amenity levels can vary from virtually none to the elaborate. Housing mobility, particularly with owner-constructed dwellings, is usually very low and consequently the markets for rental or sale of owner-constructed houses are limited and there is very little movement between the two. In principle the compilation of a house price index is the same for owner-constructed housing as for third party constructed housing, but the measurement problems are, at the least, different and are generally more difficult.⁽¹⁵⁾

9.49 The above complications mean that formal records will rarely be kept of the cost of building the new dwelling or of upgrading an old house, for example, by incorporating running water, an internal WC or additional rooms. Formal transfers of ownership sometimes do not take place, formal valuations are often not available and methods of financing can be informal through the family or may simply not be recorded or records not kept centrally. Thus in these circumstances it will not be possible to calculate mortgage interest payments (including or excluding notional interest payments to relatives), or to estimate net acquisition costs.

9.50 The lack of such basic information often means that the rental equivalence or an imputed rent approach is the only practical option for constructing a housing price index. The price indicator for imputed rents can be derived either from a readily available price series for rents, re-weighted to reflect the current composition of the stock of owner-occupier housing, which can then be applied to the rental equivalents in the base period, or from asking an expert to provide on a monthly basis the equivalent rents for a sample of houses which is representative of the owner-occupier housing stock.

9.51 In each case, stratification by type of dwelling (house or flat), location (region or area, urban or rural), plus other characteristics which will influence rent is important so that the rents data can be combined to reflect the composition of owner-occupied property. Other stratification variables may include such things as the total size of the plot, floor area and number of rooms, whether there is mains water, an internal WC and mains electricity, the material used in construction and whether the building is of traditional design. The price statistician should seek the advice of an expert active in the field of renting domestic property, such as a housing corporation,

⁽¹⁴⁾ The stratification or mix adjustment method was discussed in Chapter 4. In the example for the Dutch town of "A", many cells were indeed empty. A "matched-model" approach was suggested to cope with this problem.

⁽¹⁵⁾ In particular, the important price determining characteristics of the structure can be quite different for a developing country than for a developed country. In a developed country, there is perhaps less variation in the type of construction and the materials used whereas the quality of shanties could differ more markedly. Also land title may be missing in many instances in developing countries which again can create problems for mix adjustment and hedonic regression techniques for adjusting housing quality.

to ascertain the most important rent-determining characteristics and should bear in mind the need to keep these to a manageable number. Weights information can be derived from the latest Housing Census or Census of Population and Housing. In practice this information may not be up-to-date due to the change in the owner-occupied housing stock which can occur in the time period between censuses. Where this is the case special surveys may need to be conducted or, particularly in urban areas including townships, use made of planning applications to update the latest census.

9.52 But the measurement problems can be significant. In summary, traditional or informal dwellings are generally built by family members or other unpaid labour. The walls can be made of less durable materials such as dried clay, bamboo or latticework and the roofs can be made from reeds, straw or palm fronds or corrugated iron. The dwellings may or may not have electricity or piped water in the dwelling, let alone other facilities. Traditional dwellings are generally located in rural areas. Some associated complications when attempting to include the owner-occupier housing costs in a consumer prices index are:

- Many such dwellings are located in or very near to large cities, such as shanty-towns. These dwellings may be rented or owner-occupied and it may be difficult to obtain details of ownership. Conducting surveys can be problematic.
- There are many such dwellings in rural areas that may be built with family labour on family or unregistered land or land in “common” ownership.

In these circumstances, the concept of “ownership” becomes a grey area. Thus the definition of owner-occupied housing and what a family actually own is subject to debate and even when there is an agreed upon definition, even basic records of the number of such owner-occupied housing may not exist let alone details of the dwellings.

9.53 Relevant characteristics for the computation of a price index, that are encountered in traditional and other dwellings in the informal market include:

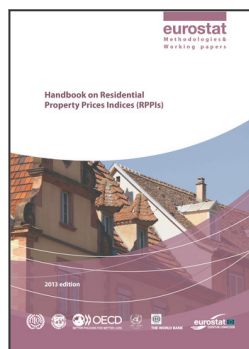
- Electricity supply. This will often be electricity supplied by a generating or distribution company. However, electricity may also be generated by the household itself, e.g. from a diesel generator or wind power, or may be taken illegally from the distributor.
- Running water. This may be piped into the dwelling itself or the dwelling takes water from a communal standpipe or well.
- A private or communal toilet, which may be either a water-flushing WC-type or a chemical toilet.

In addition there is, as with any home the issue of living space, recorded in terms of number of rooms, m², or both. For this there need to be relevant definitions. In particular, definitions of usable floor space (the floor area of the living room, kitchen, hall, bathroom and all adjoining rooms minus the wall thickness and door and window recesses and excluding e.g. stairs) and of the number of rooms (e.g. to whether to include or exclude hall-ways) are required.

9.54 Finally, even if information on the characteristics of these dwellings is available there may not be an “equivalent” rental unit to value the services of an owner-occupied unit. Thus the *indirect* measurement of prices may not be possible. In this situation, statisticians can put a system in place to measure input prices (construction costs) and then use this information to construct a user cost measure of the housing services as a proxy for the prices of the housing services consumed.⁽¹⁶⁾ For own-account consumption, the System of National Accounts 1993 (SNA 1993) recognises that it may only be practicable to measure input prices.

The issues discussed above are considered in the case study on the compilation of residential property price indices in South Africa, which can be found in Chapter 10.

⁽¹⁶⁾ See Blades (2009) for additional material on constructing these user costs for traditional housing in developing countries.



From:
Handbook on Residential Property Price Indices

Access the complete publication at:
<https://doi.org/10.1787/9789264197183-en>

Please cite this chapter as:

Fenwick, David (2013), "Data Sources", in OECD, *et al.*, *Handbook on Residential Property Price Indices*, Eurostat, Luxembourg.

DOI: <https://doi.org/10.1787/9789264197183-11-en>

This work is published under the responsibility of the Secretary-General of the OECD. The opinions expressed and arguments employed herein do not necessarily reflect the official views of OECD member countries.

This document and any map included herein are without prejudice to the status of or sovereignty over any territory, to the delimitation of international frontiers and boundaries and to the name of any territory, city or area.

You can copy, download or print OECD content for your own use, and you can include excerpts from OECD publications, databases and multimedia products in your own documents, presentations, blogs, websites and teaching materials, provided that suitable acknowledgment of OECD as source and copyright owner is given. All requests for public or commercial use and translation rights should be submitted to rights@oecd.org. Requests for permission to photocopy portions of this material for public or commercial use shall be addressed directly to the Copyright Clearance Center (CCC) at info@copyright.com or the Centre français d'exploitation du droit de copie (CFC) at contact@cfcopies.com.