Developing a local house price index: The case of Aberdeen, Scotland

Anthony Owusu-Ansah, Deborah Roberts, Rainer Schulz, and Martin Wersing*

January 2013

Abstract

The paper reports on the development of a local house price index for the Aberdeen housing market. It highlights the decisions that need to be made in the process of developing such an index and the influence of institutional context on these decisions. The basis of the index is a hedonic model re-estimated quarterly using a rolling window approach. Index values are generated using Laspeyres (base-period) weights and defining a standard property as semi-detached with median structural characteristics. The paper describes the additional information reported to users (alongside the index values) including changes over different periods, price levels and the volume of transactions. Analysis of results support the development of local house price indices and confirm that house price dynamics in the Aberdeen housing market area are distinct from those estimated for broader geographical areas.

Keywords: Hedonic regression; local house price index; housing markets

JEL Classification: R21, R31

*Owusu-Ansah, Schulz, Wersing: Centre for Real Estate Research, University of Aberdeen Business School, Edward Wright Building, Dunbar Street, Aberdeen AB24 3QY, United Kingdom. Roberts: Centre for Real Estate Research, University of Aberdeen Business School, and Social, Economic & Geographical Science Group, James Hutton Institute, Cragiebuckler, Aberdeen AB15 8QM, United Kingdom.

1 Introduction

People like to talk about house prices at dinner parties but, at the end of the meal, there are often as many opinions as there are guests. This might be due to a lack of information on actual house transactions. But a deeper problem is that houses are heterogeneous. Knowing the transaction price of a detached house in London does not provide much information on the price of a similar house in Newcastle. Similarly, knowing the price of a flat in a Georgian building in Edinburgh will not be of much help in estimating the price of a new-built flat in Bristol. The problem diminishes if transaction prices are available from the same local area but does not disappear due to the variability of housing types within any particular area. Meaningful discussion of house prices therefore requires that price information controls for quality and location differences.

A large literature exists on the methodology of computing quality-adjusted house price information using hedonic regression, see for a survey Hill (2012). Two obvious examples of hedonic-based indices in the UK context are the Halifax and Nationwide house price indices which are produced on a monthly basis at national level and quarterly basis at regional level. While being highly valued, the spatial scale at which both the Halifax and Nationwide indices are available is limiting for many users. In particular, as several authors have argued, the most appropriate scale for house information is the local level as this is the spatial scale at which information is required by planners, developers, sellers and buyers and their agents (Costello and Watkins 2002, Goh et al. 2012). While several academic studies have employed hedonic methods to produce price information at the sub-national level (e.g. Meen (1999); McDonald and Taylor (1993)) and the local level (e.g. Iacoviello and Ortalo-Magné (2003); Jones et al. (2004); Jones and Leishman (2006); Wilson et al. (2011)), this has largely been done on a one-off basis, to address particular research questions.

Against this background, this paper reports on the development of a new quality-controlled local house price index for Aberdeen, North East Scotland. As far as the authors are aware, the regular publication of a quality-controlled local house price index places Aberdeen in a unique position within the UK. The index has come about as a result of collaboration between Aberdeen Solicitors Property Centre (ASPC), Aberdeen City Council (ACC) and researchers based in the Centre for Real Estate Research (CRER), University of Aberdeen. This index will be updated on a quarterly basis.

The paper explains the institutional context and theoretical basis of the index and relates this to the numerous decisions that need to be made in order to provide price information on an ongoing basis for a range of beneficiaries. Such decisions include, for example, the choice of weights to construct the index, the selection of characteristics that reflect a "standard house" in the locality, and the period over which each hedonic regressions is estimated. In this respect, the paper contributes to an important practical, but currently small literature on the implementation of hedonic house price indices (see for example, Gouriéroux and Lafferrère (2009)).

Section 2 of the paper provides background context including an overview of the nature of the Aberdeen economy, the institutional context and the house price information that was previously made available to the public. Section 3 formally presents the hedonic regression method used to compute the index while Section 4 describes the data used to estimate the hedonic model. Section 5 presents results from the hedonic model which are subsequently used to construct the index values. It then shows how the local index compares to price indices calculated at a broader geographical scale (Scotland and UK) and to a local price index based on the average price. It also describes the other information reported to users including the volume of transactions and price levels for "standard" properties in different locations within the Aberdeen case study area. Section 6 concludes.

2 Background

The index is constructed for an area which covers both Aberdeen City Council's administrative area and the (mainly rural) Aberdeenshire Council area. In 2011, 220,420 people were estimated to live in the Aberdeen City Council area, 247,600 in Aberdeenshire with the number of households in each local authority area 103,843 and 104,724 respectively (General Register Office for Scotland 2012). There has been positive net migration in both local authorities over the last decade with increases in the Aberdeenshire area particularly noticeable, in part due to a process of counter-urbanisation (Liu and Roberts 2012).

The North Sea oil and gas industry is the main driver of the Aberdeen economy, with the total employment supported by the industry estimated to be in excess of 137,000, or 60% of all employment in the area (Aberdeen City and Shire 2011). Unemployment rates in Aberdeen City and Aberdeenshire are low compared to other regions of Scotland while the 2009 Gross Value Added per head was 43.7% higher than the UK average (Office for National Statistics 2011). The strong presence of oil and gas related activity cushions the economy from macroeconomic shocks. Concomitantly the local housing market has displayed different dynamics to housing markets elsewhere in Scotland or at the UK level (Wilson et al. 2011). There is thus a *prima facie* case for providing robust local house price information for the area.

Most residential transactions throughout Scotland are facilitated by solicitors and in several cities solicitors have grouped together to set up centres to assist them with marketing houses in their area. The Aberdeen Solicitors' Property Centre (ASPC) was the very first such centre to be established in the UK and is the model upon which other centres in Scotland have been based. While it is difficult to assess the accurately market coverage, the ASPC claims to cover 95% of all properties for sale in North East Scotland with the most obvious exclusions being new properties marketed directly by developers and larger rural properties which tend to get marketed by specialist rural surveying companies.

The centre provides a single node through which prospective house buyers can receive information about all houses on the market and the sellers can receive a better level of service than an individual solicitor could provide. This is most visible in the form of the Solicitor's Property Centre, located in the city of Aberdeen, where prospective buyers and sellers can visit and receive information on houses on offer. The Centre also runs a web-based platform, which provides information on houses listed for sale. As discussed further in Section 4, information on current and past listings are stored electronically (including various structural characteristics of properties and geo-references), and the final transaction price is recorded and provided to members for valuation purposes.

ASPC data has, since 2005, generated quarterly basis house price information for dissemination to the public. However, this information has been based on average prices of the properties transacted during a particular quarter, without formally controlling for the characteristics of those properties. The full information content of the ASPC data has thus not been fully exploited. The development of the hedonic quality-controlled index better utilises the data and, as a result, provides more reliable price information for ASPC members, other surveyors, local policy makers and planners both in Aberdeen City and Aberdeenshire Councils, banks and building societies and the general public. From a broader perspective, the index improves the information base available in a highly heterogeneous market and this, in turn, improves market efficiency.

3 Formal presentation of the hedonic approach

As is well documented, a number of alternative approaches for calculating quality-controlled indices exist including the hedonic approach, repeat sales methods and hybrid methods. In this case, the ability to apply the repeat sales and hence hybrid approach was restricted due to difficulties identifying individual flats within a building from the information provided by the ASPC. As a result, the number of repeat sales that could be identified in the data set was limited. In addition, the results form a comparison of each method, similar to that conducted by Goh et al. (2012), suggested the hedonic approach provided better price estimates.[1] The decision to adopt the hedonic approach was also consistent with the desire to show price changes for a standard property as explained further below.

3.1 Implementing the hedonic regression

The pure hedonic approach starts with the assumption that individual house prices contain information on the general price trend in an area, but are also influenced by the characteristics of a particular house. In particular, it uses regression techniques to separate the information embedded in house prices into a market-wide price trend and the value of house-specific characteristics.

To explain the approach formally, we start with the assumption that the transaction price P of a house is the product of the market value $V(\mathbf{x})$ and unusual circumstances U

$$P = V(\mathbf{x})U \; . \tag{1}$$

The market value depends on house characteristics and the general market conditions in the transaction period. This information is observed by us and is collated in the row vector $\mathbf{x}$. Unusual circumstances will occur when a buyer or a seller exploits superior knowledge about general market conditions. We expect, on average, no systematic knowledge advantage of buyers or sellers, which implies $\mathrm{E}[U] = 1$. It follows that the expected price of a house with characteristics $\mathbf{x}$ equals its market value, $\mathrm{E}[P|\mathbf{x}] = V(\mathbf{x})$. In instances of knowledge advantage, the realization of U will differ from 1 and the house will change hands above or below the market value. It is ruled out, however, that a house comes for free, which implies that realizations of U are always strictly positive. We assume further that U is independent of $\mathbf{x}$. This implies that knowing $\mathbf{x}$ *ex ante* does not help predicting the realization of U.

[1] Full details are provided in Owusu-Ansah (2012).

We specify the market value with the parametric function

$$V(\mathbf{x}) = \exp\{\mathbf{d}\boldsymbol{\gamma} + \mathbf{z}\boldsymbol{\beta}\} \; , \tag{2}$$

where $\mathbf{x} \equiv (\mathbf{d}, \mathbf{z})$. The elements of the row vector $\mathbf{d}$ include a constant and indicate the period in which the house has been sold. For instance, if we cover four quarters and the house was transacted in the third, then $\mathbf{d} = [1\,0\,1\,0]$ and it follows that $\mathbf{d}\boldsymbol{\gamma} = \gamma_1 + \gamma_3$. If the house was transacted in the fourth quarter instead, then $\mathbf{d} = [1\,0\,0\,1]$ and $\mathbf{d}\boldsymbol{\gamma} = \gamma_1 + \gamma_4$. The elements of the row vector $\mathbf{z}$ give information about house-specific characteristics. If we take a house with unchanged characteristics $\mathbf{z}$, then the change of its market value between quarters three and four is simply $\exp\{\gamma_4 - \gamma_3\}$.

Taking logs of Eqs. 1 and 2, we obtain the hedonic regression equation

$$p = \mathbf{d}\boldsymbol{\gamma} + \mathbf{z}\boldsymbol{\beta} + u \tag{3}$$

with $p \equiv \ln P$ and $u \equiv \ln U - \mathrm{E}[\ln U]$.[2] We have $\mathrm{E}[u] = 0$ and $\mathrm{E}[\mathbf{x}u] = \mathbf{0}$ because $\mathbf{x}$ and u are statistically independent. Eq. 3 shows that the price of a house depends on the general market trend $\mathbf{d}\boldsymbol{\gamma}$ and the observed house characteristics $\mathbf{z}$ valued with the $\boldsymbol{\beta}$ coefficients. A structural interpretation of the $\boldsymbol{\beta}$ coefficients as marginal valuations is tempting, but requires that the supply of characteristics is inelastic. This might be plausible in the short but is less so in the long run. Consistent estimation of marginal valuations also requires that we observe all characteristics that impact on the market value.

The coefficients $\boldsymbol{\gamma}$ and $\boldsymbol{\beta}$ in the linear model in Eq. 3 can be estimated with ordinary least squares. The discussion above indicated that the elements in $\boldsymbol{\beta}$ are not necessarily constant through time, because preferences for characteristics may change and progress in building technologies may lead to supply of improved building characteristics. To account for this possibility, we implement a rolling window

[2]To ease exposition, we subsume the constant term $\mathrm{E}[\ln U]$ into the coefficient γ_1.

estimation approach using the regression equation

$$p_i = \mathbf{d}_i \boldsymbol{\gamma}_t + \mathbf{z}_i \boldsymbol{\beta}_t + u_i \; . \tag{4}$$

Here, p_i is the log price of house i, the vector $\mathbf{z}_i$ collects observed characteristics and u_i is unsystematic noise. Eq. 4 shows that price variation over time is driven by changes in the general price trend, $\boldsymbol{\gamma}_t$, and by changes of the market valuation of characteristics, $\boldsymbol{\beta}_t$. The coefficient vectors $\boldsymbol{\gamma}_t$ and $\boldsymbol{\beta}_t$ have time subscripts to indicate that they may vary over time. We fit the regression in Eq. 4 using information on the houses transacted between the current t period and periods back to $t - w$. An estimation sample covers therefore $1 + w$ periods, with $n_{\tau,t}$ observations in period $\tau \in \{1, \ldots, 1 + w\}$ and a total number of observations

$$N_t \equiv \sum_{\tau=1}^{1+w} n_{\tau,t} \; .$$

Once new data become available, the estimation window shifts forward by one quarter.

Eq. 4 can be written compactly as

$$\mathbf{p}_t = \mathbf{D}_t \boldsymbol{\gamma}_t + \mathbf{Z}_t \boldsymbol{\beta}_t + \mathbf{u}_t \; . \tag{5}$$

The $(N_t \times 1)$ vector $\mathbf{p}_t$ contains the log prices of the houses transacted in the estimation sample that ends in period t. The $(N_t \times (1 + w))$ matrix $\mathbf{D}_t$ contains the constant term and indicates the periods in which the houses were transacted. The $N_t \times K$ matrix $\mathbf{Z}_t$ contains the observed characteristics of the houses and the $(N_t \times 1)$ vector $\mathbf{u}_t$ contains the individual error terms. Applying ordinary least squares to Eq. 5 leads to

$$\begin{bmatrix} \widehat{\boldsymbol{\gamma}}_t \\ \widehat{\boldsymbol{\beta}}_t \end{bmatrix} = \begin{bmatrix} \mathbf{D}_t'\mathbf{D}_t & \mathbf{D}_t'\mathbf{Z}_t \\ \mathbf{Z}_t'\mathbf{D}_t & \mathbf{Z}_t'\mathbf{Z}_t \end{bmatrix}^{-1} \begin{bmatrix} \mathbf{D}_t' \\ \mathbf{Z}_t' \end{bmatrix} \mathbf{p}_t \; . \tag{6}$$

The estimators for the unobserved coefficients $\boldsymbol{\gamma}_t$ and $\boldsymbol{\beta}_t$ are functions of the observed characteristics, $\mathbf{Z}_t$, prices, $\mathbf{p}_t$, and the matrix considering the constant and the time dummies, $\mathbf{D}_t$. A slight reformulation of Eq. 5 is instructive for the interpretation of the estimator for $\boldsymbol{\gamma}_t$. To do so, we exploit that

$$\mathbf{D}_t = \mathbf{A}_t \mathbf{T} \tag{7}$$

with the $N_t \times (1+w)$ matrix

$$\mathbf{A}_t \equiv \begin{bmatrix} \mathbf{1}_{n_{1,t}} & \mathbf{0}_{n_{1,t}} & \mathbf{0}_{n_{1,t}} & \dots & \mathbf{0}_{n_{1,t}} \\ \mathbf{0}_{n_{2,t}} & \mathbf{1}_{n_{2,t}} & \mathbf{0}_{n_{2,t}} & \dots & \mathbf{0}_{n_{2,t}} \\ \mathbf{0}_{n_{3,t}} & \mathbf{0}_{n_{3,t}} & \mathbf{1}_{n_{3,t}} & \dots & \mathbf{0}_{n_{3,t}} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \mathbf{0}_{n_{1+w,t}} & \mathbf{0}_{n_{1+w,t}} & \mathbf{0}_{n_{1+w,t}} & \dots & \mathbf{1}_{n_{1+w,t}} \end{bmatrix}$$

and the $(1+w) \times (1+w)$ matrix

$$\mathbf{T} \equiv \begin{bmatrix} 1 & \mathbf{0}'_w \\ \mathbf{1}_w & \mathbf{I}_w \end{bmatrix} .$$

The unit (zero) column vector $\mathbf{1}_n$ ($\mathbf{0}_n$) has n elements and $\mathbf{I}_w$ is a $w \times w$ identity matrix. Using Eq. 7 in Eq. 5 gives

$$\mathbf{p}_t = \mathbf{A}_t \boldsymbol{\theta}_t + \mathbf{Z}_t \boldsymbol{\beta}_t + \mathbf{u}_t \tag{8}$$

with

$$\boldsymbol{\theta}_t \equiv \mathbf{T}\boldsymbol{\gamma}_t = \begin{bmatrix} \gamma_{1,t} \\ \gamma_{1,t} + \gamma_{2,t} \\ \dots \\ \gamma_{1,t} + \gamma_{1+w,t} \end{bmatrix} . \tag{9}$$

If we apply ordinary least squares to the regression in Eq. 8, we obtain a system similar to Eq. 6. Bringing the system into normal equations form allows us to write

$$\widehat{\boldsymbol{\theta}}_t = (\mathbf{A}'_t \mathbf{A}_t)^{-1} \mathbf{A}'_t \mathbf{p}_t - (\mathbf{A}'_t \mathbf{A}_t)^{-1} \mathbf{A}'_t \mathbf{Z}_t \widehat{\boldsymbol{\beta}}_t . \tag{10}$$

The $(1+w) \times N_t$ matrix

$$(\mathbf{A}_t'\mathbf{A}_t)^{-1}\mathbf{A}_t' = \begin{bmatrix} \frac{1}{n_{1,t}}\mathbf{1}_{n_{1,t}}' & \mathbf{0}_{n_{2,t}}' & \mathbf{0}_{n_{3,t}}' & \cdots & \mathbf{0}_{n_{1+w,t}}' \\ \mathbf{0}_{n_{1,t}}' & \frac{1}{n_{2,t}}\mathbf{1}_{n_{2,t}}' & \mathbf{0}_{n_{3,t}}' & \cdots & \mathbf{0}_{n_{1+w,t}}' \\ \mathbf{0}_{n_{1,t}}' & \mathbf{0}_{n_{2,t}}' & \frac{1}{n_{3,t}}\mathbf{1}_{n_{3,t}}' & \cdots & \mathbf{0}_{n_{1+w,t}}' \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \mathbf{0}_{n_{1,t}}' & \mathbf{0}_{n_{2,t}}' & \mathbf{0}_{n_{3,t}}' & \cdots & \frac{1}{n_{1+w,t}}\mathbf{1}_{n_{1+w,t}}' \end{bmatrix}$$

pre-multiplied simply computes the averages of observed prices and characteristics for each period in the estimation sample, respectively. Eq. 10 has then a straightforward interpretation: the estimated time dummy coefficient for period τ in the estimation sample is the average log price for the period corrected for the impact of observed characteristics evaluated at the current implicit prices

$$\widehat{\theta}_\tau = \overline{p}_\tau - \overline{\mathbf{z}}_\tau \widehat{\boldsymbol{\beta}}_t \; . \tag{11}$$

It follows from the above decomposition that an index based on the quarterly average log price is a good measure of the general price trend if the average characteristics of the transacted houses remain the same over time *and* if the implicit prices stay constant.[3] If either of them or both change over time, then the average price is a distorted measure of the general price trend.[4]

What happens if the market value function is not log linear, as assumed in Eq. 3, or if more house characteristics are relevant for market values than observed by us? In this case, the ordinary least squares estimator in Eq. 6 is the best linear predictor under squared error loss given the observed house characteristics (Cameron and Trivedi 2005, 4.2.3). It is therefore better than the average log price, which is a linear predictor too, but one that ignores house characteristics.

[3] Such an index has the form $\exp\{\overline{p}_\tau\}$ and equals the geometric average of prices.

[4] We treat the case where both components of $\overline{\mathbf{z}}_\tau \widehat{\boldsymbol{\beta}}_t$ vary in such a way that the product stays constant over time as irrelevant.

3.2 Constant-quality index series

Once the estimates in Eq. 6 have been computed for a estimation window ending in period t, a Laspeyres quality-controlled price index is constructed.[5] Let $\mathbf{d}_c$ denote a $1 \times (1+w)$ vector that has 1 as first and last element and 0 elsewhere. Multiplying this vector with the estimate $\widehat{\boldsymbol{\gamma}}_t$ picks the first and the last coefficient. The first coefficient is the overall constant of the estimated regression and the last coefficient measures the relative level of the general house price trend in quarter t. The sum of both coefficients is the estimate for $\theta_{1+w,t}$, see Eq. 10. Let $\mathbf{z}_c$ denote the $1 \times K$ vector with the characteristics of a specified standard house. Multiplying this vector with the estimate $\widehat{\boldsymbol{\beta}}_t$ gives the log value of the characteristics.

The quality-controlled price index for period t with base period $t = 0$ is

$$I_t \quad = \quad \frac{\exp\{\mathbf{d}_c\widehat{\boldsymbol{\gamma}}_t + \mathbf{z}_c\widehat{\boldsymbol{\beta}}_t\}}{\exp\{\mathbf{d}_c\widehat{\boldsymbol{\gamma}}_0 + \mathbf{z}_c\widehat{\boldsymbol{\beta}}_0\}}\ . \tag{12}$$

Eq. 12 is our Laspeyres base period weighted index. The index relates the estimated market value for the standard house in period t to the market value of the identical house in the base period $t = 0$.

4 Data and model implementation

This section describes the data and explains the modelling decisions that had to be made to implement the Aberdeen quality-controlled house price index. The latter involves the choice of variables included in the hedonic regression, its functional form, as well as the length of the estimating sample window. Moreover, housing sub-areas within Aberdeen and Aberdeenshire and the characteristics of a "standard house" within these areas had to be chosen in order to provide as disaggregated price information as possible.

[5] There are several potential problems associated with base weighted indices as discussed in Diewert (2007). However it is the most easily interpreted form of index and this was considered in getting the new index accepted by stakeholders.

4.1 Description of data and choice of variables

The data are provided by the ASPC and consists of all transactions of residential properties in Aberdeen City and Aberdeenshire between 2000Q1 to 2012Q2. After cleaning the raw data, 76,731 transactions remained.[6] Each observation has information on the transaction price, address, postcode area, location coordinates, sale date, dwelling type (whether a flat, semi-detached or detached property), whether it is a new or old building, numbers of floors, public rooms, bedrooms, bathrooms and whether the property has a garage and/or garden. Details on the data cleaning are given in Appendix A.

The data has shortcomings, because not all characteristics relevant for market values are observed. First, the literature suggests the size of a property is useful in explaining its price and is usually represented in models by lot size or floor area (Sirmans et al. 2005). In the absence of such information, the number of rooms is used as a proxy for property size (Leishman 2001). Second, apart from a dummy variable distinguishing new builds from vintage houses, no indication of the age of the property is given in the data. This restricts the ability of the model to explore the effect of depreciation on house prices.

Table 1 gives summary statistics on the transaction price and seven structural house characteristics known for each transaction.

[Table 1 about here]

As indicated by the standard deviation, house prices show substantial variation. This is in line with the substantial variation of structural house characteristics, such as number of floors and rooms. In the hedonic regression, all of the variables are treated as discrete and converted into dummy variables, see also Section 4.3 below. The decision to treat the number of public rooms, bedrooms and bathrooms as

[6]In total, about 9 percent of transactions were dropped from the raw data.

dummy rather than continuous variables was made to avoid imposing a log-linear relationship in prices associated with these structural characteristics.

4.2 Choice of housing sub-areas

Rather than have a single index for the whole North East of Scotland, feedback from users suggested that information on separate housing sub-areas would be more useful. Thus decisions had to be made on which areas to distinguish when providing the house price information. In negotiation with the ASPC, five separate sub-areas were selected as the basis for presenting the results: Aberdeen City and Suburbs, Ellon, Inverurie, Stonehaven, and the remaining "Other countryside" area. Of these, Aberdeen City and Suburb area is by far the largest in terms of number of transactions (see Table 1). The three separately identified towns, Ellon, Inverurie, and Stonehaven, were selected on the basis of being large but distinct urban centres in Aberdeenshire with clear connections to the city of Aberdeen, while the "Other countryside" category captures all of the remaining localities in Aberdeenshire apart from i) the three separately distinguished towns and ii) the localities classed as being within the Aberdeen suburb area.

Having selected the five sub-areas, decisions had to be made in relation to their boundaries. After discussions with users, the geographical boundaries for the Aberdeen City and Suburb area were defined to be identical to those used by the ASPC itself when marketing properties. Table 2 gives an overview of the resulting localities of the Aberdeen City and Suburbs area, as well as localities in Aberdeenshire.

[Table 2 about here.]

While the delineation of housing sub-areas by these administrative areas counters recent academic arguments suggesting empirical identification of sub-market boundaries (Leishman 2009, Goh et al. 2012), it reflects the institutional context within which the index was constructed.

4.3 Specification of hedonic regression model

Fitting the hedonic regression model given in Eq. 4 to the data requires to specify its functional form, as well as to decide on the number of periods each estimation sample should cover.

As mentioned above, we treat all structural house characteristics as discrete and convert them into dummy variables. To further increase the flexibility of the hedonic model, the impact of the structural characteristics are allowed to vary by the three dwelling types (detached, semi-detached and flat) *within* each of the five housing sub-areas under consideration. In addition to structural house characteristics, we also include location coordinates in the regression to capture differences in the value of neighborhood amenities.

With respect to the length of the estimation sample, we use 21 quarters ($w = 20$) to fit each hedonic regression model. This choice ensures that the estimation sample is sufficiently large to generate statistically robust estimate of the model parameters.

The resulting specification of Eq. 4 is

$$\begin{aligned} p_i &= \gamma_{1,t} + \sum_{\tau=2}^{21} I_{\tau,t}(i)\,\gamma_{\tau,t} + \sum_{j=1}^{14} I_{j,t}(i)\,\beta_{j,0,t} + \sum_{j=1}^{15}\sum_{k=1}^{12} I_{j,t}(i)\,z_{i,k}\beta_{j,k,t} \\ &\quad + f_t(LAT_i, LON_i) + u_i\ , \end{aligned} \tag{13}$$

where $I_{\tau,t}(i)$ is a binary indicator variable that takes the value one if house i was sold in period τ of the current estimation sample. This sample runs from $t-20$ ($\tau = 1$) to the current period t ($\tau = 21$). Otherwise, the indicator takes the value zero. $I_{j,t}(i)$ is a binary indicator variable that takes the value one if house i in the estimation sample belongs to dwelling type sub-area combination j ($j = 1, \ldots, 15$) and zero otherwise. The reference category for the dwelling type sub-area combinations is a flat in Aberdeen City and Suburbs ($j = 15$). House characteristics excluded from vector $\mathbf{z}$ are: one bedroom, one public room, one floor, one bathroom, no garage, and no garden. The excluded characteristics serve as reference categories. $f_t(\cdot)$ is a function of location coordinates that is parameterized as

$$f_t\left(LAT_i, LON_i\right) = \beta_{1,f} LAT_i + \beta_{2,f} LON_i + \beta_{3,f} LAT_i^2 + \beta_{4,f} LON_i^2 + \beta_{5,f} LAT_i \times LON_i \ ,$$

where LAT_i and LON_i represent the location coordinates of house i and $\beta_{k,f}$ are unknown coefficients that have to be estimated.

4.4 Definition of standard property

To construct index series for each of the five housing sub-area the characteristics of a "standard" property have to be specified. While this may appear trivial, the choice will clearly influence the dynamics of the index relative to other indices, as well as the accessibility of the index to users and is therefore important. For the Aberdeen City and Suburbs index, the standard property is defined on the basis of the median characteristics of a semi-detached dwelling as reported in Table 3.

[Table 3 about here]

In particular, it is defined to be semi-detached with three bedrooms, two floors, two public rooms, one bathroom, no garage, and a garden. While a flat is the most commonly occurring property type in the region and thus perhaps a more obvious basis for defining a standard property, the choice of a semi-detached property ensured that the price levels reported when the index was publicly launched in 2012Q2 remained close to the average house price reported by the ASPC previously.

The standard property is treated as identical across all five sub-areas despite differences in the nature of housing stock in each one. This decision was driven by the desire to simplify comparisons in price movements across areas. The characteristics of the standard property will be reviewed and, if necessary, revised on a 10 year basis to ensure changes in building stock over time are recognised.

5 Results

5.1 Hedonic regression results

Table 4 summarizes ordinary least squares estimates of Eq. 13 for the initial estimation window of the index, which covers the period 2000Q1 to 2005Q1.[7] For the ease of exposition, we report only coefficients on house characteristics for flats, detached and semi-detached houses that are located in Aberdeen City and Suburbs. Standard errors are calculated with a robust variance estimator to allow for heteroscedastic error terms in the regression.

[Table 4 about here]

The overall fit of the regression is remarkably well with a $\bar{R}^2$ of 0.737. Moreover, the estimated coefficients are statistically significant at the usual levels most of the times. They also have plausible signs and magnitudes. An increase in the number of rooms (public rooms, bedrooms and bathrooms) has generally a positive and economically significant impact on house prices. An increase from one bedroom to two bedroom, for instance, increases the price of a flat (semi-detached, detached house) in Aberdeen City by approximately 44.8 (30.4, 38.5) percent. Similarly, a flat (semi-detached, detached house) with two public rooms is priced approximately 34.8 (22.6, 17.6) percent higher as compared to the same property type with only one public room. Properties with a garage and/or a garden also sell with a premium. A flat (semi-detached, detached house) with a garage, for example, demands a price increase of about 18.4 (18.4, 8.6) percent higher as compared to a property with no garage.

[7] Estimates for subsequent periods are of similar quality as the results discussed here.

5.2 House price indices

Based on the findings from the hedonic model and the characteristics of the standard property (as defined in section 4.4), five constant-quality house price indices can be constructed for the North East of Scotland which are shown in Figure 1. The period of observation is 2005Q1 to 2012Q2 and the base period of each index is the first quarter of 2005 (2005Q1=100).

[Figure 1 about here.]

The index number rises from 100 to 161.9 in 2012Q2 for Aberdeen City and Suburbs, 165.3 for Ellon, 165.3 for Inverurie, 159.3 for Stonehaven and 158.8 for the countryside area. As expected the five indices follow a similar pattern of price change over time and share the same turning points, peaking in early 2007. However some differences are apparent with the Other Countryside area consistently being above the levels of the other areas and differences between the series appearing to be larger following the fall in prices in 2007.

To further explore the importance of having local house price indices, Figure 2 compares the Aberdeen City and Suburbs index with the Scottish regional and the UK national indices constructed by Halifax (Fleming and Nellis 2013) and the retail price index (RPI) for the United Kingdom from the Office for National Statistics.

[Figure 2 about here.]

It is obvious from Figure 2 that house price levels in Aberdeen have, over the period analyzed, been comparatively high and volatile as compared to house prices at the Scottish or UK level. In particular, compared to a 61.9% rise in house prices in Aberdeen City and Suburbs, the increase for Scotland over the whole period was only 6.6% while at UK level prices declined by 0.9%. The finding thus supports the need for the provision of local house price information to better understand regional housing markets.

Figure 3 compares the constant-quality price index to the (geometric) average index.[8]

[Figure 3 about here.]

As expected, the figure shows that the average indices are more volatile than the quality adjusted indices. This confirms that an average index which does not control for the heterogeneity of properties may give an unreliable picture of price changes over time.

5.3 Other information disseminated to beneficiaries

To complement the quarterly house price indices, the following more disaggregated information is published in a quarterly house price bulletin. First, the results from the hedonic model are used to calculate annual house price changes, and annualized house price changes over a five-year holding period for each of the five housing sub-areas. All price changes are published alongside changes in the UK Retail Price Index (RPI) estimated by the Office for National Statistics, the Halifax Scottish regional house price index, and the Halifax UK national house price index.[9] Table 5 presents the quarterly, annual, and 5-year annualized price changes for 2012Q2.

[Table 5 about here]

Clearly, the price dynamics for the five Aberdeen housing areas show the unique dynamic of the housing market in North East Scotland as compared to the aggregate price behavior in Scotland and the UK, respectively.

[8] This index is computed as

$$\frac{\exp\{\overline{p_t}\}}{\exp\{\overline{p}_0\}}$$

[9] Annualized price changes for the UK and Scotland are calculated using unadjusted index values from Halifax.

Second, following feedback from the ASPC and Aberdeen City Council, the price level of three different "standard" properties in each of the five areas are provided. Again, the characteristics of the different standard properties are based on the median characteristics reported in Table 3 and are treated as identical across all five sub-areas.

[Table 6 about here]

Table 6 reports the quality-adjusted prices in Aberdeen City and the surrounding housing areas for 2012Q1 and 2012Q2. As expected the standard house in the Aberdeen City and Suburb area sells with a significant price premium as compared to the rural housing sub-areas in Aberdeenshire.

Finally, based on feedback from surveyors, information on the volume of transactions is provided. Figure 4 shows the quarterly volume of house transaction for the period 2005Q1 to 2012Q2.

[Figure 4 about here]

The average number of transactions occurring in per quarter between 2000Q1 to 2012Q2 is just over 1,500. Trade volumes, however, follow a seasonable pattern with sales peaking in the third quarter and being lowest in the first quarter of a given year. Moreover, trade volumes are positively correlated with the growth rate of house prices.

6 Conclusion

Reliable information on house prices not only provides something to talk about at dinner parties, it is also important information for developers, lenders, sellers and buyers and their agents. The process of implementing a local index requires several decisions to be made which influence not only the results but also the utility of the information to users. It is argued that while there has been considerable attention in the housing literature to the most appropriate choice of modelling approach to index construction, very little attention has been given to other more detailed decisions required if a house price index is to be produced on an ongoing basis to a range of different users. These include the frequency of publication of the index, how to deal with changes in preferences over time or changes in the nature of building stock, and how to select the characteristics of a "standard" house within the area. In addition, decisions need to be made on the range of information to provide to users.

Against this background, the paper reports on the development of a new local house price index for the Aberdeen housing market, explaining the context, theoretical nature of the index, underlying data, and the nature of results that are provided to potential beneficiaries. The paper provides a detailed explanation of the process and decisions made in the development of the index. In this case, five separate Laspeyres index series are calculated based on the results from a hedonic model estimated each quarter using a 21-period rolling-window approach. The results are compared to indices for Scotland and the rest of the UK. The results strongly support the production of local house price information. Moreover, comparison with an index based solely on transaction values confirms the importance of allowing for the heterogeneity of housing stock in an area.

The institutional context influenced the nature of the index in several ways. First, the desire of users to know the level as well as changes in house prices influenced the choice of modelling approach, while their desire to maintain a degree of consistency with the level of previously published (average-based) house price information influenced the bundle of characteristics selected for the âœstandard propertyâ. Discussions with stakeholders rather than theoretical reasoning also influenced the selection of boundaries for the sub-areas and more generally, the content and presentation of findings included in a quarterly bulletin.

There are several refinements of the index and related information that could be implemented in the future. First, we currently model a general price trend for each of the five areas including Aberdeen City and Suburbs. It is conceivable, however, that price trends differ between market segments within the city, such as flats and detached houses. Associated with this, areas of high-priced housing and affordability issues in Aberdeen City co-exist alongside depressed areas with low housing demand. Further analysis of the location variables in the model would provide valuable information to policy makers and planners within the City area. Second, we report predicted house prices, which correspond to the average price we expect for a house with given characteristics: it would be useful to report quantiles prices as well as this average figure. Third, price changes are only one part of the return the owner of a house receives. If the owner also occupies the house, a notional rent should also be included in order to calculate the total return from being a house owner. This would place the information on a comparable basis to that available to commercial real estate investors. Notional rents could, potentially be assessed by using market rents for rental houses with the same characteristics. The ASPC has recently started to provide information on residential lease data and we hope that in time this could be used to provide fuller information on the benefits of residential investment.

Acknowledgements

We are grateful for helpful comments by Verity Watson. The usual disclaimer applies. Financial support from the University of Aberdeen through the PhD studentship for Anthony Owusu-Ansah is gratefully acknowledged.

A Data cleaning

The data cleaning proceeds in the following steps:

1. Remove transactions that include more than one property with no indication of whether the properties involved were homogenous.

2. Remove transactions with no information on one or more of the seven structural house characteristics included in the hedonic regression.

3. Remove observations with extreme prices whose log price is more than three standard deviations away from the yearly average log price.

References

Aberdeen City and Shire: 2011, *Housing Need and Demand Assessment Update*, Aberdeen.

Cameron, A. C. and Trivedi, P. K.: 2005, *Microeconometrics. Methods and Applications*, Cambridge University Press, New York.

Costello, G. and Watkins, C.: 2002, Towards a system of local house price indices, *Housing Studies* **17**(6), 857–873.

Diewert, W. E.: 2007, The paris oecd-imf workshop on real estate price indexes: Conclusions and future directions, *Ubc departmental archives*, UBC Department of Economics.

Fleming, M. C. and Nellis, J. G.: 2013, The Halifax house price index. technical details, *Technical report*, Halifax.

General Register Office for Scotland: 2012, *Mid-2011 Population Estimates Scotland. Population Estimates by Sex, Age and Administrative Area*, Edinburgh.

Goh, Y. M., Costello, G. and Schwann, G.: 2012, Accuracy and robustness of house price index methods, *Housing Studies* **27**(5), 643–666.

Gouriéroux, C. and Lafferrère, A.: 2009, Managing hedonic price indexes: The French experience, *Journal of Housing Economics* **18**(3), 206–213.

Hill, R. J.: 2012, House price indexes for residential housing: A survey, evaluation and taxonomy, *Journal of Economic Surveys* (forthcoming, doi: 10.1111/j.1467-6419.2012.00731.x).

Iacoviello, M. and Ortalo-Magné, F.: 2003, Hedging housing risk in London, *Journal of Real Estate Finance and Economics* **27**(2), 191–209.

Jones, C. and Leishman, C.: 2006, Spatial dynamics of the housing market: An interurban perspective, *Urban Studies* **43**(7), 1041–1059.

Jones, C., Leishman, C. and Watkins, C.: 2004, Exploring linkages between housing submarkets: Theory, evidence and policy implications, *Housing Studies* **19**(2), 269–283.

Leishman, C.: 2001, House building and product differentiation : An hedonic price approach, *Journal of Housing and the Built Environment* **16**, 131–152.

Leishman, C.: 2009, Spatial change and the structure of urban housing sub-markets, *Housing Studies* **24**(5), 563–585.

Liu, N. and Roberts, D.: 2012, Counter urbanisation, planning and house prices: An analysis of the Aberdeen housing market area, *Town Planning Review* (forthcoming).

McDonald, R. and Taylor, M. P.: 1993, Regional house prices in Britain: Long-run relationships and short-run dynamics, *Scottish Journal of Political Economy* **40**(1), 43–55.

Meen, G.: 1999, Regional house prices in Britain: Long-run relationships and short-run dynamics, *Housing Studies* **14**(6), 733–753.

Office for National Statistics: 2011, *Regional Gross Value Added (Income Approach)*, Newport.

Owusu-Ansah, A.: 2012, Measuring and understanding the house price dynamics of the Aberdeen housing market. PhD Thesis University of Aberdeen.

Sirmans, G. S., Macpherson, D. A. and Zietz, E. N.: 2005, The composition of hedonic pricing models, *Journal of Real Estate Literature* **13**(1), 3–43.

Wilson, P., White, M., Dunse, N., Cheong, C. and Zurbruegg, R.: 2011, Modelling price movements in housing micro markets: Identifying long-term components in local housing market dynamics, *Urban Studies* **48**(9), 1853–1874.

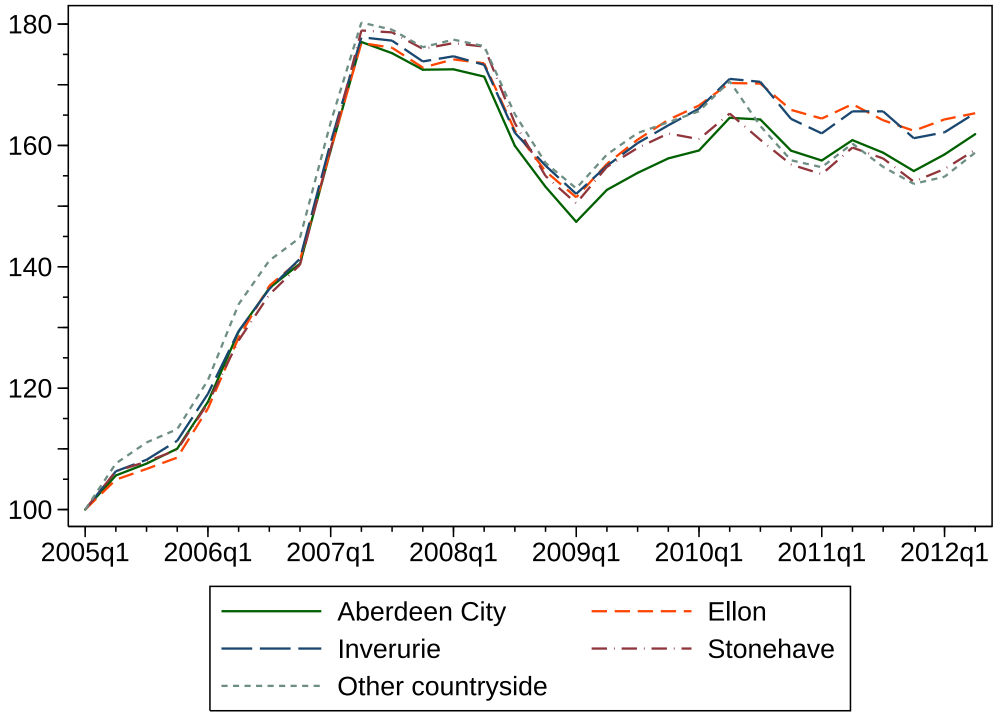

Figure 1: Price trends for the Aberdeen housing market. Figure shows constant-quality house price indices for 5 housing sub-areas in Aberdeen City and Aberdeenshire. Base period is first quarter of 2005 (2005Q1=100).

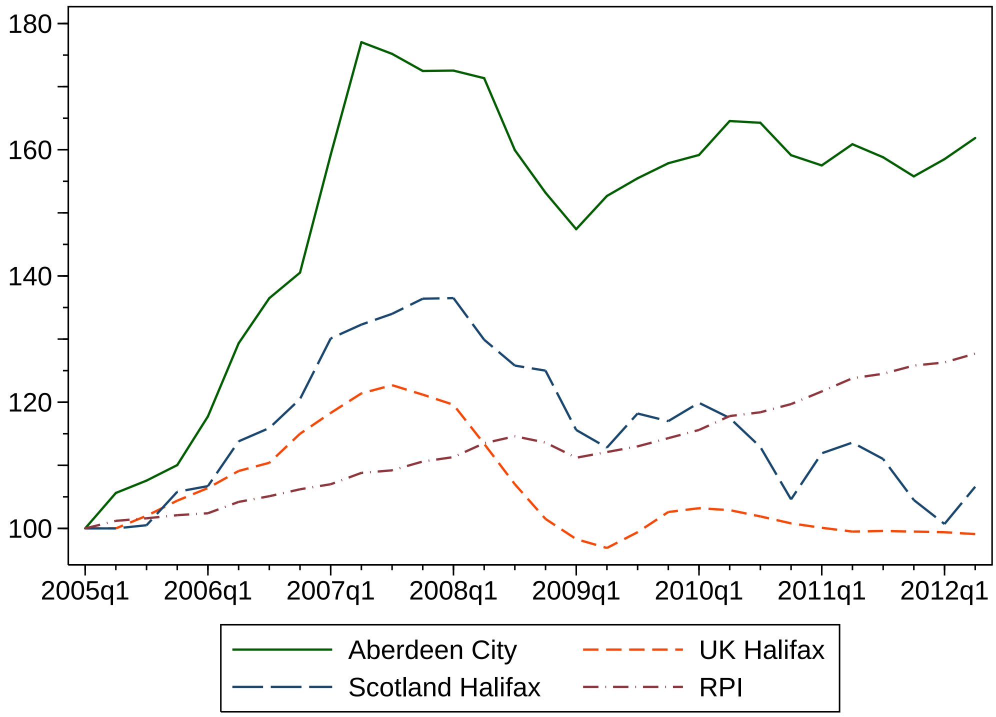

Figure 2: Aberdeen house price trend and the Scottish regional and UK national price trends. Figure shows constant-quality house price index for Aberdeen City and Suburbs, the Halifax house price index for Scotland and the UK, respectively, and the UK retail price index. Base period is first quarter of 2005 (2005Q1=100).

Figure 3: The quality-adjusted and average house price trends. Figure shows the constant-quality house price index and average house price indice for Aberdeen City and Suburbs. Base period is first quarter of 2005 (2005Q1=100).

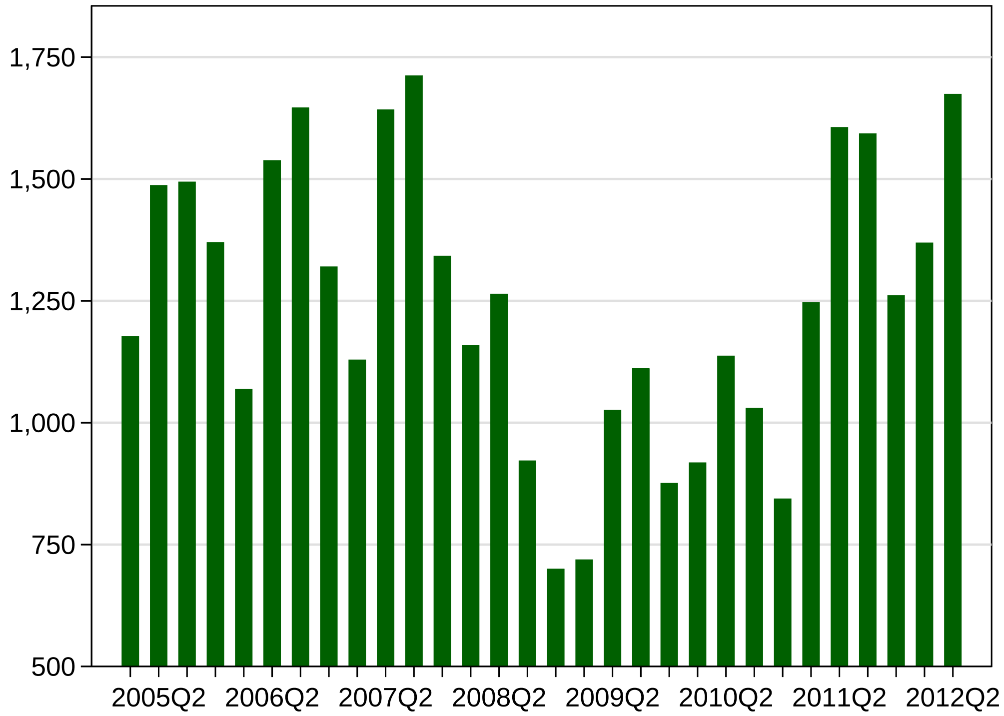

Figure 4: Volume of house transactions in Aberdeen City and Aberdeenshire areas from 2005Q1- 2012Q2.

Table 1: Summary statistics for all transacted houses 2000Q1 to 2012Q2. Number of observations is 76,731. House price is in real (year 2000) Pounds Sterling.

	Mean	Median	Std. Dev.
House price			
Number of floors	1.503	1	0.536
Number of public rooms	1.551	1	0.790
Number of bedrooms	2.514	2	1.068
Number of bathrooms	1.035	1	0.183
Has a			
Garage	0.387		
Garden	0.684		
Dwelling Type			
Detached	0.272		
Semi-detached	0.307		
Flat	0.420		
Housing sub-area			
Aberdeen city	0.735		
Ellon	0.022		
Inverurie	0.028		
Stonehave	0.029		
Other country areas	0.186		

Table 2: Composition of housing sub-aeras. Table summarizes localities of Aberdeen City and suburbs area and localities in Aberdeenshire. *Locality forms separate housing sub-area and is excluded from the Other Countryside area.

Aberdeen City and Suburbs	Countryside
Aberdeen city centre	Aboyne
Balmedie	Alford
Banchory-Devenick	Ballater/Braemar
Bankhead/Bucksburn/Stoneywood	Banchory
Bieldside	Banff
Blackburn	Brechin/Montrose
Blairs	Buckie
Bridge of Don	Collieston/Newburgh
Cove Bay/Findon	Cruden Bay
Cults/Pitfodels	Cullen
Danestone	Drumoak/Durris
Dyce	*Ellon
Elrick/Skene/Westhill	Fraserburgh
Kingswells	Fyvie/New Deer/Turrif
Maryculter	Gardenstown/Macduff
Milltimber	Huntly/Keith
Muchalls/Newtonhill	Insch
Nigg	*Inverurie
Newmachar	Inverbervie/Johnshaven
Peterculter	Kemnay
Portlethen	Kintore
Potterton	Laurencekirk
	Lumphanan
	Methlick/Tarves
	Monymusk
	Oldmeldrum/Pitmedden/Udny
	Peterhead
	Portsoy
	*Stonehaven
	Torphins

Table 3: Summary statistics for the various dwelling types: 2000Q1 to 2004Q4. Number of observations are 14,441 for flats, 9,764 for semi-detached and 8,701 for detached.

	Number of				Has a	
	Bedrooms	Public rooms	Floors	Bathrooms	Garage	Garden
Flat						
Median	2	1	1	1	0	0
Mean	1.711	1.134	1.052	1.012	0.071	0.294
Semi-detached						
Median	3	2	2	1	0	1
Mean	2.609	1.582	1.894	1.038	0.462	0.977
Detached						
Median	4	2	2	1	1	1
Mean	3.430	2.200	1.716	1.095	0.814	0.981

Table 4: Hedonic regression, 2000Q1-2005Q1. Table reports OLS estimates of coefficients in Eq. 13 for houses located in Aberdeen City and suburbs. Constant, time dummies, and location coordinates are not reported. Number of observations is 34,275. Heteroscedasticity robust standard errors are reported in parenthesis. *** significant at 1%-level ** significant at 5%-level * significant at 10%-level.

	Flat		Semi-detached		Detached	
Type			0.069	[0.029]**	0.431	[0.076]***
2 Bedrooms	0.448	[0.006]***	0.304	[0.010]***	0.385	[0.058]***
3 Bedrooms	0.575	[0.014]***	0.447	[0.011]***	0.554	[0.057]***
4 Bedrooms	0.731	[0.030]***	0.729	[0.020]***	0.758	[0.058]***
5 Bedrooms	0.741	[0.057]***	0.888	[0.035]***	0.974	[0.060]***
2 Public rooms	0.348	[0.011]***	0.226	[0.007]***	0.176	[0.012]***
3 Public rooms	0.458	[0.026]***	0.451	[0.014]***	0.339	[0.014]***
4 Public rooms	0.710	[0.053]***	0.635	[0.026]***	0.528	[0.022]***
2 Floors	0.168	[0.021]***	-0.075	[0.008]***	-0.043	[0.011]***
3 Floor	0.585	[0.215]***	0.175	[0.027]***	0.139	[0.034]***
2 Bathrooms	0.090	[0.031]***	-0.024	[0.019]	0.061	[0.018]***
Garage	0.184	[0.012]***	0.184	[0.007]***	0.086	[0.016]***
Garden	0.026	[0.007]***	0.129	[0.029]***	-0.063	[0.062]
$\bar{R}^2$	0.737					

Table 5: Rate of price changes in UK, Aberdeen and the surrounding areas for 2012Q2. All figures are in percent.

	Quarterly price change 2012Q1 to 2012Q2	Annual price change: 2011Q2 to 2012Q2	Annualized price change over 5 year holding period: 2007Q2 to 2012Q2
Aberdeen City and suburbs	2.1	0.6	-1.8
Ellon	0.6	-0.9	-1.3
Inverurie	2.0	-0.2	-1.4
Stonehaven	2.0	-0.2	-2.3
Other country areas	2.6	-0.9	-2.5
Scotland (Halifax)	5.9	-5.4	-4.3
UK (Halifax)	-0.3	-0.5	-4.0
Retail Price Index (ONS)	1.1	3.1	3.3

Table 6: Quality-adjusted house prices in Aberdeen and the surrounding areas for 2012Q2. The house prices are in Pounds Sterling. The housing characteristics used to estimate the house prices are two bedrooms, one floor, one public room, one bathroom, no garage and no garden for flat; three bedrooms, two floors, two public rooms, one bathroom, no garage and a garden for semi-detached; four bedrooms, two floors, two public rooms, one bathroom, a garden and a garage for detached.

Property type	Period	Aberdeen City and suburbs	Ellon	Inverurie	Stonehaven	Other countryside
Flat	2012Q2	145,222	126,440	127,805	134,562	123,191
	2012Q1	142,913	123,675	125,569	133,643	122,085
Semi-detached	2012Q2	193,334	161,320	185,482	179,603	174,160
	2012Q1	189,333	160,330	181,919	176,056	169,795
Detached	2012Q2	310,678	246,305	270,827	282,754	292,981
	2012Q1	304,515	241,945	267,106	279,536	286,491