A hedonic house price index in continuous time

Sofie R. Waltl
October 2015

Department of Economics
Department of Public Economics
University of Graz

An electronic version of the paper may be downloaded from the RePEc website: http://ideas.repec.org/s/grz/wpaper.html
A hedonic house price index in continuous time

Sofie R. Waltl

University of Graz, Austria
Universitätstraße 15/F4
8010 Graz, Austria
Tel.: +43 316 380 3441
sofie.waltl@uni-graz.at

October 21, 2015

Abstract

Purpose: This paper develops a methodology to measure movements in housing markets accurately and timely by constructing a continuously estimated house price index.

Design/methodology/approach: The continuous index, which is extracted from an additive model that includes the temporal as well as the locational effect as smooth functions, can be interpreted as an extension of the classical hedonic time-dummy method. The methodology is applied to housing sales from Sydney, Australia, between 2001 and 2011 and compared to three types of discrete indexes.

Findings: Discrete indexes turn out to approach the continuously estimated index with decreasing period lengths but eventually become wiggly and unreliable due to too few observations per period. The continuous index, in contrast, is stable, has some favourable robustness properties and is more objective in several ways.

Originality/value: The resulting index tracks movements in the housing market precisely and in “real-time” and is hence suited for monitoring and assessing housing markets. Since turbulence in housing markets is often a harbinger of financial crises, such monitoring tools support policy makers and investors in tailoring their decisions and reactions. Additionally, the index can be evaluated arbitrarily frequently and therefore is well suited for use in property derivatives.

Keywords: Residential property, House price indexes, Hedonic indexes, Continuous indexes, Additive Models, Housing market analysis

Paper type: Research paper
1 Introduction

Sharply rising house prices are usually a harbinger of financial crises. The recent global financial crisis starting in 2007 is a prominent example of this (see for instance Claessens et al., 2010). Having a tool at hand that tracks movements in housing markets in “real-time” can serve as an early warning system. The proposed index combines a high level of precision in detecting turning points with high accuracy in measuring the true changes in price levels. A continuous house price index that can be updated every day helps policy makers improving their reactions. Besides that, high-frequency indexes may be used in derivative markets. The prominent S&P/Case-Shiller Home Price Indices only report monthly changes. Derivatives based on a more frequent house price index could be created using the methodology presented in this paper adding more options to shape an investor’s exposure to the housing market without going through the extensive process of buying or selling a property.

The proposed index can be evaluated at any point in time within the period of observation. Discrete indexes, however, deliver only one average index number per period. Such indexes detect movements in the housing market only between periods but not within. Peaks and troughs within a period may be averaged out. The main argument for aggregating house sales within a period is, that prices do not change rapidly. But still, it is important to detect changes once they occur regardless of the speed of change to avoid undetected and wrongly placed turning points, misleading or weakened trends and wrong price levels. In a continuous estimation nothing like a period has to be specified, which drops a source of subjectivity as indexes are quite sensitive towards these choices. Also the starting points of a period are subjective model specifications. For instance, letting yearly periods start on the first of January or the first of July may imply very distinct indexes. However, the choice of period lengths and starting points is not driven by such considerations. These kinds of shortcomings are overcome by switching from discrete to continuous indexes. To a certain extent, median and time-dummy indexes approach the continuously estimated index with decreasing period lengths demonstrating the continuous index’s preciseness. Time is a continuous variable and should be treated as such.

The main advantage of continuous indexes is that current transactions are linked to preceding transactions. Schwann (1998) used a time-series-based approach to reach this goal. In Schwann’s approach the intercept is the only time-dependent component. In reality, also other parameters and especially the valuation of location may vary over time. A method that extends Schwann’s approach with this regard is proposed by Francke and Vos (2004). Despite these examples, house price indexes generally rely on regression techniques rather than time-series approaches. The index I propose here follows this tradition and also allows for flexible shadow prices. The Divisia price index formula is another approach to construct time-continuous indexes, which to my knowledge has not been applied in a housing context thus far (see Divisia, 1926; Hulten, 1973).
The model used to construct the continuous index is an extension of the hedonic time-dummy method. The use of hedonic methods to construct house price indexes is recommended in the *Handbook on Residential Property Prices Indices* jointly developed by the European Union, International Labour Organization, International Monetary Fund, Organisation for Economic Co-operation and Development, United Nations Economic Commission for Europe and The World Bank (see de Haan and Diewert, 2013). The hedonic equation is estimated using *Additive Models* (Hastie and Tibshirani, 1993) that allow non-linear estimation of predictor variables. Many studies have pointed out that including house characteristics in a non-linear way is crucial (see Shimizu et al., 2014, for details and further references). Here I go a step further and estimate also the temporal effect non-linearly.

The paper is structured as follows: Section 2 explains the construction of the continuous index, an extension allowing flexible shadow prices and derives standard errors. Section 3 describes the dataset used to derive the empirical results presented in section 4. Section 5 discusses robustness properties and section 6 concludes.

## 2 A time-continuous index

### 2.1 The standard approach

A time-dummy index results from the following hedonic equation:

$$\log P = D\delta + X\beta + \varepsilon, \quad (1)$$

where $P$ denotes the vector of transaction prices, $X$ the matrix of structural and locational house characteristics and $\beta$ its associated (unknown) shadow prices. $D$ is a matrix of dummy variables indicating the period of transaction, $\delta$ denotes period-specific intercepts and $\varepsilon \sim iid \mathcal{N}(0, \sigma^2)$. The series of estimated period-specific intercepts $\hat{\delta}_t, t = 1, \ldots, T$, build the basis of the house price index as they describe changes in the housing market net of effects driven by house characteristics. The semi-log form of the hedonic equation implies that the estimates $\hat{\delta}_t$ are on a logarithmic scale and have to be back-transformed. This is usually done by just taking the exponent, $\exp(\hat{\delta}_t)$.

$$\hat{\delta}_t = \exp(\hat{\delta}_t + \bar{x}\hat{\beta}).$$

Normalizing yields$$\frac{\hat{P}_t}{\hat{P}_{t^*}} = \exp(\hat{\delta}_t + \bar{x}\hat{\beta}) = \frac{\exp(\hat{\delta}_t)}{\exp(\hat{\delta}_{t^*})},$$
implying the same index as before. This second approach to extract an index is more convenient when dealing with the continuous version presented below.

From probability theory it follows that \( \exp(\hat{\delta}_t) \) is a biased estimator, which was already pointed out by Kennedy (1981). Model (1) entails

\[
\hat{\delta}_t \sim \mathcal{N}\left(\delta, \sigma^2((\tilde{X}^\top \tilde{X})^{-1})_{tt}\right),
\]

where \( \tilde{X} = (D, X) \) denotes the full design matrix. \( \exp(\hat{\delta}_t) \) follows a log-normal distribution with

\[
\mathbb{E}\left[ \exp(\hat{\delta}_t) \right] = \exp\left(\delta + \frac{1}{2} \sigma^2((\tilde{X}^\top \tilde{X})^{-1})_{tt}\right).
\]

This suggests estimating \( \exp(\hat{\delta}_t) \) by the mean-unbiased estimator

\[
\exp\left(\hat{\delta}_t - \frac{1}{2} \hat{\sigma}^2((\tilde{X}^\top \tilde{X})^{-1})_{tt}\right).
\]

From the properties of a log-normal distribution, it is on the other hand evident that

\[
\text{Median}\left[ \exp(\hat{\delta}_t) \right] = \exp(\delta).
\]

Hence, \( \exp(\hat{\delta}_t) \) is a median-unbiased estimator! Estimating the development of median rather than mean house prices is advantageous as the median is less sensitive towards outliers and therefore a more stable indicator for general movements in the housing market. Hedonic indexes that track changes in median prices can be directly compared to the simple median and stratification approach. Furthermore, the mean-unbiased estimator relies on the estimated variance \( \hat{\sigma}^2 \). The quality of the estimator \( \hat{\sigma}^2 \) largely depends on the accuracy of the underlying distributional assumption, i.e., on the assumption of log-normally distributed house prices. The median estimator is less influenced by this assumption. Empirical analyses further show that the magnitude of the bias term, \( \frac{1}{2} \hat{\sigma}^2((\tilde{X}^\top \tilde{X})^{-1})_{tt} \) is very small and therefore in practice no great differences between mean-unbiased and median-unbiased estimators are to be expected (see for instance Syed et al., 2008). Also my own calculations for the dataset I am using in this paper yield bias terms almost identical to zero.

Following these arguments, I will ignore the bias term in the discrete as well as in the continuous case.

To gain a continuous index, a continuous time scale constructed from the exact transaction dates is needed:

\[
\text{TIME}_i = \text{YEAR}_i + \frac{\text{MONTH}_i - 1 + \frac{\text{DAY}_i - 1}{30}}{12}.
\]

Let \( \text{TIME} = (\text{TIME}_1, \cdots, \text{TIME}_n)^\top \) and \( n \) be the number of observations. Model (1) changes to

\[
\log P = f(\text{TIME}) + X\beta + \varepsilon,
\]
where \( f(\cdot) \) is a smooth function that is estimated non-parametrically yielding an Additive Model (see Appendix A for more details).

Following the ideas of Hill and Scholz (2014) locational effects are measured smoothly on a two-dimensional grid spanned by longitudes and latitudes, \( f_2(\text{LONG, LAT}) \). A transition from region dummies to a smoothly estimated function is a gain in precision as discrete measurement of the continuous locational effect is substituted by continuous measurement of the former. This logic is equivalent to the transition from a discrete time measure to a continuous measure of time.

Summarizing other variables describing the dwellings’ physical characteristics in the matrix \( X \) finally yields the model

\[
\log P = f_1(\text{TIME}) + f_2(\text{LONG, LAT}) + X\beta + \varepsilon.
\]

An index can be constructed in the same manner as in the case of time-dummy models. Given a combination of house characteristics \( \bar{x} \) and a specific location \((\text{LONG, LAT})\), the model is evaluated at these characteristics and a (narrowly spaced) sequence of points in time \((\text{TIME}_1, \ldots, \text{TIME}_t, \ldots, \text{TIME}_T)\) within the period of observation. A very convenient way is to evaluate the index on a daily basis.\(^2\) The resulting predicted house prices are back-transformed and normalized,

\[
\frac{\hat{P}_t}{\hat{P}_{t'}} = \frac{\exp \left( \hat{f}_1(\text{TIME}_t) + \hat{f}_2(\text{LONG, LAT}) + \bar{x}\hat{\beta} \right)}{\exp \left( \hat{f}_1(\text{TIME}_{t'}) + \hat{f}_2(\text{LONG, LAT}) + \bar{x}\hat{\beta} \right)} = \frac{\exp \left( \hat{f}_1(\text{TIME}_t) \right)}{\exp \left( \hat{f}_1(\text{TIME}_{t'}) \right)}.
\]

Although the index is only evaluated at a discrete set of points, it is still continuous as the model estimates the time effect continuously and allows to calculate an index value at any arbitrary point in time within the period of observation as a consequence of the smoothing algorithm applied. The index does not report averaged values over a period in the evaluation process but precise numbers for every point in time, which is the most important advantage of the continuous index over any discrete index.

2.2 Accounting for changing shadow prices

The standard approach assumes that valuation of house characteristics and location does not change over time, which is a very restrictive assumption.

Next to the temporal effect, there are three types of predictor variables entering the hedonic model: categorical and continuous variables and the two-dimensional locational effect. Each type is treated differently as described in the following. To keep notation simple, assume that just one categorical variable \( X^{\text{cat}} \) with \( L \) levels and one continuous variable \( X^{\text{cont}} \) enter the model, i.e.,

\[
\log P = \beta_0 + f_1(\text{TIME}) + f_2(\text{LONG, LAT}) + \beta^{\text{cont}} X^{\text{cont}} + \sum_{l=2}^{L} \beta_l^{\text{cat}} \mathbf{1}(l)(X^{\text{cat}}) + \varepsilon,
\]
where

\[ I_{\{l\}}(X^{\text{cat}}) = \begin{cases} 1, & X^{\text{cat}} = l, \\ 0, & X^{\text{cat}} \neq l. \end{cases} \]

Both, categorical and continuous covariates are allowed to vary over time by introducing interactions. Categorical variables are directly interacted with the smoothly estimated time effect. Thus,

\[ \sum_{l=2}^{L} \beta_{\text{cat}}^{l} I_{\{l\}}(X^{\text{cat}}) \]

is replaced by

\[ \sum_{l=2}^{L} \beta_{\text{cat}}^{l} I_{\{l\}}(X^{\text{cat}}) + \sum_{l=2}^{L} f_{\text{cat}}^{l}(\text{TIME}|X^{\text{cat}} = l) I_{\{l\}}(X^{\text{cat}}). \]

\( f_{\text{cat}}^{l}(\text{TIME}|X^{\text{cat}} = l) \) is a smooth function which considers in the estimation process only observations satisfying \( X^{\text{cat}} = l \). So, additionally to the main effect associated with the \( l \)th level \( \beta_{\text{cat}}^{l} \) a second, time-dependent effect \( f_{\text{cat}}^{l}(\text{TIME}|X^{\text{cat}} = l) \) enters the model. The total effect of level \( l \) is then just the sum of the two, \( \beta_{\text{cat}}^{l} + f_{\text{cat}}^{l}(\text{TIME}|X^{\text{cat}} = l) \). Per construction, \( f_{\text{cat}}^{l}(\text{TIME}|X^{\text{cat}} = l) \) is centred around zero and measures the deviation from the main effect \( \beta_{\text{cat}}^{l} \).

Continuous variables could directly be interacted with a continuous time function, which would yield a two-dimensional effect \( f(\text{TIME}, X^{\text{cont}}) \). This would – particularly in case of many continuous variables – dramatically increase the complexity of the model and complicate the interpretation. It is more convenient to categorize a continuous variable appropriately and add interactions between this categorized variable and the continuous time effect. Let \( X^{\text{cont}} \) be split into \( C \) categories yielding \( X^{\text{cont}|C} \). Then,

\[ \beta_{\text{cont}} X^{\text{cont}} \]

is replaced by

\[ \beta_{\text{cont}} X^{\text{cont}} + \sum_{c=2}^{C} f_{\text{cont}}^{c}(\text{TIME}|X^{\text{cont}|C} = c) I_{\{c\}}(X^{\text{cont}|C}). \]

Also the locational effect could directly be interacted with the smooth time effect yielding a three-dimensional effect \( f(\text{LONG, LAT, TIME}) \). For the same reasons as before, one would prefer to avoid such complex model structures and hence categorize one of the two continuous effects. Categorizing the locational effect means switching back from using longitudes and latitudes to the much less precise approach of relying on postcode or region dummies. \( f(\text{LONG, LAT}) \) allows locational effects to vary continuously over space and any kind of discretization would lower the level of precision of this most important effect in housing markets. Hence, categorizing the time effect is to be preferred in this case. Ideally, the smooth function is updated regularly and frequently (e.g., monthly) by interacting \( f(\text{LONG, LAT}) \) with period dummies. This is, however, subject to two practical limitations: First, if the chosen period is too
short, there are not enough observations to gain reliable estimates. Second, updating the non-parametric term frequently is a computational burden. Such an index is still a good benchmark to analyse the importance of time-dependent locational effects. In the following I will call this index the benchmark index.

A more convenient method is based on overlapping periods. The basic idea is to estimate $J$ models that allow $f(LONG, LAT)$ to update at rare intervals. These models all rely on the same length of updating intervals but differ in the starting points of the intervals. From every model on predicts daily prices $P_{t}^{\text{model}_j}$, which are averaged by using a geometric mean, $P_{t}^{\text{avg}} = \left( \prod_{j=1}^{J} P_{t}^{\text{model}_j} \right)^{1/J}$. After normalizing, $P_{t}^{\text{avg}}/P_{t}^{*}$, the averaged index results.

The benchmark index has a major drawback: For discrete indexes the selection of period lengths and starting points influences the resulting index and brings in an additional source of subjectivity. Choosing appropriate updating intervals is subject to the same considerations. The averaged index tries to account for this subjectivity by averaging over many possible starting points of the updating interval.

2.3 Standard errors for the continuous index

The continuous index is constructed by evaluating the continuously estimated time effect at arbitrarily many points of time within the period of observation. Let $M$ be the prediction matrix$^3$, i.e., the matrix by which the estimated coefficients are multiplied to get the (untransformed) index:

$$\hat{Y} = M\hat{\beta}.$$  

$M$ consists of constant values for all house characteristics but different values for the time variable (for instance a list of all days within the period of observation). $\hat{Y}$ is then a vector that describes the time effect evaluated at a daily basis, $\hat{Y} = (\hat{Y}_t)_{t=1}^T$ for $T$ days. For additive models approximately

$$\hat{\beta} \sim \mathcal{N}(\beta, V_\beta)$$

holds, where $V_\beta$ denotes the covariance matrix$^4$ associated with the model parameters (see Wood, 2006). Therefore, $\hat{Y} \sim \mathcal{N}(M\beta, MV_\beta M^\top)$ and

$$\text{Var}[\exp(\hat{Y}_t)] = \exp \left( 2 \cdot (M\beta)_t + (MV_\beta M^\top)_{tt} \right) \left( e^{(MV_\beta M^\top)_{tt}} - 1 \right).$$

The estimated standard errors are then obtained by using plug-in estimates

$$\hat{s.e.}(\exp(\hat{Y}_t)) = \sqrt{\exp \left( 2 \cdot (M\hat{\beta})_t + (M\hat{V}_\beta M^\top)_{tt} \right) \left( e^{(M\hat{V}_\beta M^\top)_{tt}} - 1 \right)}.$$  

From these standard errors point-wise confidence intervals are calculated to measure the precision of the price index.
3 Data

I use a dataset created by Australian Property Monitors that consists of house transactions in Sydney within the period 2001 to 2011. The dataset has two main advantages: first, a very large number of observations and, second, exact longitudes and latitudes of each house sale. Its limitations are a small number of physical characteristics (land area, number of bed- and bathrooms) and missing values (46.3% of the observations are incomplete). The latter does not induce a problem since observations are incomplete at random (see Appendix B). To avoid misleading results, I restrict the analysis to houses with transaction prices within the range 100,000 and 4 million Australian dollars, having at most six bed- or bathrooms and a land area of less than 5,000 square meters. Table 1 reports summary statistics for all complete (and refilled as described in the following) observations.

<table>
<thead>
<tr>
<th></th>
<th>PRICE</th>
<th>AREA</th>
<th>BED</th>
<th>BATH</th>
</tr>
</thead>
<tbody>
<tr>
<td>Minimum</td>
<td>100,000</td>
<td>100.0</td>
<td>1</td>
<td>101,587</td>
</tr>
<tr>
<td>1st Quartile</td>
<td>368,000</td>
<td>462.0</td>
<td>2</td>
<td>95,632</td>
</tr>
<tr>
<td>Median</td>
<td>520,000</td>
<td>587.0</td>
<td>3</td>
<td>31,992</td>
</tr>
<tr>
<td>Mean</td>
<td>654,611</td>
<td>624.3</td>
<td>4</td>
<td>4,489</td>
</tr>
<tr>
<td>3rd Quartile</td>
<td>765,000</td>
<td>716.0</td>
<td>5</td>
<td>762</td>
</tr>
<tr>
<td>Maximum</td>
<td>4,000,000</td>
<td>4,996.0</td>
<td>6</td>
<td>138</td>
</tr>
<tr>
<td>Longitudes:</td>
<td>(150.6000, 151.3396)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Latitudes:</td>
<td>(−34.19957, −33.42726)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Number of observations:</td>
<td>435,295</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Number of complete observations:</td>
<td>233,724</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Number of complete observations after reconstruction:</td>
<td>280,471</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Table 1: Summary statistics.

The dataset also includes a unique identifier for each house that can be used to detect houses that changed hands several times. These observations can be used to reconstruct at least some missing values. If for instance a house appears twice in the dataset and the number of bedrooms is available both times but the number of bathrooms is available only for the second transaction, then the observed number of bathrooms of the first transaction is assumed to be also the number of bathrooms of the house at the date of the second transaction. To avoid wrong reconstructions, gaps are refilled if and only if all of the following constraints are met:

1. **Constancy constraint:** Available numbers of bedrooms (or bathrooms) are constant, i.e., there are no contradicting recordings for the same house.

2. **Time constraint:** The time span between two transactions is greater than six months:

   \[ \text{TIME}_{\text{diff}} = \text{TIME}_2 - \text{TIME}_1 > 0.5 \text{ years}. \]
3. **Price growth constraint:** The average annual price growth is less than 25%:

\[
\left( \frac{\text{PRICE}_2}{\text{PRICE}_1} \right)^{1/\text{TIMEdiff}} - 1 < 25%.
\]

The reconstruction algorithm reduces the share of incomplete recordings from 46.3% to 35.6%! Non-refillments occurred mainly due to a violation of the constancy constraint.

4 **Results**

4.1 **The hedonic model**

The model equation for the standard approach is given by

\[
\log P = \beta_0 + f_1(\text{TIME}) + f_2(\text{LONG}, \text{LAT}) + \sum_{i=2}^{6} \beta_i^{\text{BED}} \mathbb{I}_{(i)}(\text{BED}) \\
+ \sum_{i=2}^{6} \beta_i^{\text{BATH}} \mathbb{I}_{(i)}(\text{BATH}) + \beta^{\text{AREA}} \log(\text{AREA}) + \varepsilon. \tag{4}
\]

[Figure 1 about here.]

Estimated parameters (in parenthesis below the respective parameter) readily fulfil expectations: The higher the number of bed- or bathrooms the higher the price of a dwelling. Only the parameter to the indicator of six bedrooms is slightly lower than the one corresponding to five bedrooms. Estimated standard errors are small (all below 0.017) and all parameters are highly significant according to t-tests. The model’s explained deviance equals 89% indicating a very good model fit.

The locational effect is shown in Figure 1. Dark colors indicate a low and light colors a high price level. As expected, higher price levels are estimated in the inner city around the Opera House (marked by the white triangle) and along the coastline. The locational effect is highly significant according to an F-test.

Figure 1 also shows the continuous index together with point-wise standard error bands. The index detects a rapid price increase until 2004, a second peak at the end of 2007 and decreasing prices thereafter. From 2009 on prices rose again before they stabilized at an all-times high level from the beginning of 2010 on. Next to this general patterns the index very precisely measures small movements around the long-term trends. Standard error bands are in general narrow indicating stable estimates.

4.2 **Time-dummy and median indexes**

[Figure 2 about here.]

[Figure 3 about here.]
Discrete indexes are per construction constant over one period, but usually are interpolated from one period to the next when plotted, i.e., the index is plotted as a continuous function although it is in fact a step function. Figure 2 (a, b) shows interpolated indexes together with more accurate step function indexes demonstrating that interpolations perform a shift of the index (magnitude is the period length) which yields wrongly placed turning points. Discrete indexes visualized in that way imitate a continuous index, however without overcoming the shortcomings resulting from discrete measurement.

A discrete index reports average values per period. Peaks and troughs of the true (but unknown) index might be averaged out. Longer periods lead to more pronounced averaging effects. Figure 2 (c) demonstrates this effect: The solid line shows the true price index. The estimated discrete index reports the average price level within each period (dotted lines) measuring a price change of \( \Delta p^* \). The true price change \( \Delta p \), however, is much bigger. In times of rapid price changes, this effect is very pronounced whereas in times of relatively constant prices the effect diminishes, i.e., the magnitude of the effect is not constant over time leading to another error source of discrete indexes. Users of discrete house price indexes assume that indexes report price changes of average houses. In fact, however, they report average price changes of average houses and I claim that there is one “average” too much.

Discrete indexes are sensitive towards the selection of starting points. Figure 2 (d) shows two yearly time-dummy indexes which respectively rely on the 1st of January and the 1st of July as starting points. The resulting indexes differ strongly – particularly in terms of index levels (the ranges are \([1, 1.7]\) and \([1, 1.56]\)!). The deviations result from distinct normalization periods and different magnitudes of the averaging effect. Again, a continuously estimated index is not affected by such choices and hence more objective.

Figure 3 (a) compares the continuously estimated index to various time-dummy and median indexes based on different period lengths. Median indexes seem to get more reliable when shortening periods from years to half-years or even quarters but more wiggly and hence less precise as period lengths are further decreased. The averaging effect is well seen when analysing time-dummy indexes. As prices rose sharply at the beginning of the time span, the averaging effect leads to far too low price levels for long periods such as years or half-years. The averaging effect becomes less pronounced for quarterly and monthly periods. In this sense, time-dummy indexes converge to the continuously estimated index. However, as median indexes the preciseness of time-dummy indexes declines when periods are shorten too much as a consequence of too few observations per period (see Figure 3 (b)). A continuously estimated index avoids the optimization problem of finding an accurate period length.

If the goal is to have only one index number per period, e.g., per year, it is preferable to estimate the index continuously and evaluate the model only once per year. The resulting values are precise numbers at specific points in time and not averages.
over a period. Hence, interpolating between index values makes sense. Figure 4 (a) compares the daily evaluated continuous index to its yearly evaluated counterpart (evaluation date: 1st of January). The resulting index is less precise in terms of detecting small movements but levels are in general measured accurately as there are no averaging effects. This kind of indexes are called semi-continuous in the following.

[Figure 4 about here.]

4.3 Stratification index

The Australian Bureau of Statistics (ABS) publishes house price indexes for the eight capital cities of Australia including Sydney. The quarterly based index uses a stratification approach. There are ten variables describing the mix of houses as well as socio-economic conditions of each suburb (see Australian Bureau of Statistics, 2005). Based on these variables, the ABS performs a principal-component analysis identifying sufficiently homogeneous strata (Sydney: 55 strata). Data comes from public authorities guaranteeing high reliability and comprehensiveness.

Despite the fact that the continuously estimated index presented in this paper and the ABS index differ in both, the methodology of construction and data used, they still aim to measure the same effect. Therefore, it is worth comparing them.

Figure 4 (b) shows the ABS index (Q3/2003 - Q2/2014) together with the continuously estimated index (1.1.2001-31.12.2011). To minimize the impact of averaging effects, I normalize both indexes with respect to the center of the overlapping period, i.e., Q3/2007 and October 1, 2007 respectively. In this quarter prices did not change a lot leading to a small averaging effect. The indexes coincide almost perfectly indicating that both indexes meet their goal to extract the pure quality-adjusted change in (median) house prices. In concordance with the prior findings, the quarterly based index seems to be an approximation of the continuous one. In general, both indexes indicate very similar trends and detect identical turning points.

In their information paper the ABS states that there were some user requests for a more frequent (i.e., monthly) index. However, the ABS “does not believe that the currently available data are sufficient to support the construction of a reliable monthly series.” This is surely true when using stratification methods, however, applying hedonic approaches allows to construct arbitrarily frequent (even continuous), reliable indexes.

4.4 Numerical results

This section measures deviations of indexes numerically. Indexes defined over different periods cannot be compared directly but have to be transformed first. For illustration purposes, let’s assume that a monthly index, \( \text{index}^m \), shall be compared to a quarterly index, \( \text{index}^q \). The period of observation be one year. So the monthly
(quarterly) index consists of \( n_m = 12 \) (\( n_q = 4 \)) values. The shorter index is expanded
\[
\text{index}^q = (q_1, q_2, q_3, q_4)\top \quad \sim \quad \text{index}^q = (q_1, \ldots, q_1, q_2, \ldots, q_2, q_3, \ldots, q_3, q_4, \ldots, q_4)\top
\]
yielding an index \( \text{index}^q \) of length \( n_m \) which can be compared to the longer index using the following measures:
\[
\max_{i \in \{1, \ldots, n_m\}} |\text{index}^q_i - \text{index}^m_i| \quad \text{maximum norm (MN)},
\]
\[
\sqrt{\sum_{i=1}^{n_m} (\text{index}^q_i - \text{index}^m_i)^2} \quad \text{Euclidean norm (EN)},
\]
\[
\frac{1}{n_m} \sum_{i=1}^{n_m} |\text{index}^q_i - \text{index}^m_i| \quad \text{average absolute distance (AAD)}.
\]

To compare the continuous index to discrete indexes, I evaluate the continuous index daily and transform the discrete indexes accordingly. \textit{Semi-continuous} indexes report values for specific points in time and not averages over a period, hence the correct way of comparing them to the daily evaluated index is to use linear interpolation between adjacent evaluation points. The results are given in Table 2.

<table>
<thead>
<tr>
<th></th>
<th>Years</th>
<th>Half-years</th>
<th>Quarters</th>
<th>Months</th>
</tr>
</thead>
<tbody>
<tr>
<td>Median</td>
<td>MN</td>
<td>0.318</td>
<td>0.254</td>
<td>0.254</td>
</tr>
<tr>
<td></td>
<td>EN</td>
<td>9.920</td>
<td>5.187</td>
<td>4.706</td>
</tr>
<tr>
<td></td>
<td>AAD</td>
<td>0.146</td>
<td>0.067</td>
<td>0.055</td>
</tr>
<tr>
<td>Time-dummy</td>
<td>MN</td>
<td>0.248</td>
<td>0.141</td>
<td>0.096</td>
</tr>
<tr>
<td></td>
<td>EN</td>
<td>9.667</td>
<td>4.954</td>
<td>2.859</td>
</tr>
<tr>
<td></td>
<td>AAD</td>
<td>0.146</td>
<td>0.075</td>
<td>0.043</td>
</tr>
<tr>
<td>Stratification</td>
<td>MN</td>
<td>–</td>
<td>–</td>
<td>0.074</td>
</tr>
<tr>
<td></td>
<td>EN</td>
<td>–</td>
<td>–</td>
<td>1.763</td>
</tr>
<tr>
<td></td>
<td>AAD</td>
<td>–</td>
<td>–</td>
<td>0.030</td>
</tr>
<tr>
<td>Semi-continuous</td>
<td>MN</td>
<td>0.080</td>
<td>0.053</td>
<td>0.016</td>
</tr>
<tr>
<td></td>
<td>EN</td>
<td>1.400</td>
<td>0.835</td>
<td>0.252</td>
</tr>
<tr>
<td></td>
<td>AAD</td>
<td>0.016</td>
<td>0.009</td>
<td>0.003</td>
</tr>
</tbody>
</table>

|                     | Time-dummy indexes with periods consisting of ten days: | MN 0.094, EN 2.653, AAD 0.040. |
|                     | five days: | MN 0.136, EN 2.895, AAD 0.042. |

Table 2: Numerical comparison to the continuous index.

Regardless which measure is taken, almost identical results are obtained. Median indexes perform worst. In consistency with prior findings, median indexes become more reliable when decreasing the period length from years to half-years and even quarters but become very wiggly and hence untrustworthy thereafter. Time-dummy indexes seem to converge to the continuously estimated index with decreasing period length. Reducing the period length even more, the index becomes less accurate.
similarly as for median indexes. The ABS stratification index is closer to the continuous index than comparable median or time-dummy indexes. The semi-continuous indexes outperform all other presented indexes. (They are per construction very similar to the continuous index.)

4.5 Flexible shadow prices

In a first step, only parametric terms are allowed to evolve over time ($\text{AREA}_{\text{cat}}$ denotes the categorized land area variable.):

\[
\log P = \beta_0 + f_1(\text{TIME}) + f_2(\text{LONG, LAT}) \\
+ \sum_{i=2}^{6} \beta_i^\text{BED} \mathbb{1}_i(\text{BED}) + \sum_{i=2}^{6} f_i^\text{BED}(\text{TIME}|\text{BED} = i) \mathbb{1}_i(\text{BED}_t) \\
+ \sum_{i=2}^{6} \beta_i^\text{BATH} \mathbb{1}_i(\text{BATH}) + \sum_{i=2}^{6} f_i^\text{BATH}(\text{TIME}|\text{BATH} = i) \mathbb{1}_i(\text{BATH}) \\
+ \beta^\text{AREA} \log(\text{AREA}) + \sum_{i=2}^{5} f_i^\text{AREA}(\text{TIME}|\text{AREA}_{\text{cat}} = i) \mathbb{1}_i(\text{AREA}_{\text{cat}}) + \varepsilon.
\]

The resulting index is shown as index (a) in Figure 5. The benchmark index (shown as index (d)) is obtained by interacting $f_2(\text{LONG, LAT})$ with yearly dummy variables. More frequent updates are computationally impracticable. However, the resulting index is prone to jumps as yearly updates are too rare. Relying on overlapping periods is another way to allow flexible locational effects. Therefore, I estimate six models with different updating structures (see Figure 5, bottom panel) of $f_2(\text{LONG, LAT})$ and calculate the geometric mean of the resulting indexes to gain the averaged index (index (c)). The averaged index is smoother and has almost no jumps. Index (c) and (d) indicate that the price peak in 2004 was much higher than suggested by the standard index indicating that an average locational effect over the entire period of observation is not precise enough. Effects that are explained by changes in locational variables are absorbed into the price index when $f_2(\text{LONG, LAT})$ is kept constant over time.

5 Robustness Analysis

Time fixity is one of the most important properties of house price indexes. From a theoretical point of view, the continuous index might change a bit when adding observations of a new period but not dramatically. This is due to the fact that I use a local basis approach to construct the spline that estimates the functional form of the time effect in the model. Using local basis functions instead of global ones guarantees that only the very end of an index might change due to new observations. Further minor differences in indexes constructed out of models spanning different
time periods result from the control of wiggliness through the smoothing parameter. Panel (a) in Figure 6 shows three continuously estimated indexes. The first one covers the entire time span, whereas the other two are restricted to the time spans 2001-2010 and 2001-2008 (i.e., leaving out the three years that experienced a rapid increase in house prices). It can be clearly seen that there are in fact no differences between the presented indexes demonstrating that time fixity is not an issue for these kinds of indexes.

Next, I analyse changes in the index if the data is not recorded on a daily but rather on a more infrequent basis. In practise the exact transaction date might not be available but only the month or quarter in which the transaction occurred. To construct a smooth time scale out of monthly data, I assume that all transactions occurred in the middle of the month. The time scale is then given by

$$\text{TIME}\cdot M_i = \text{YEAR}_i + \frac{\text{MONTH}_i - 1 + \frac{15 - 1}{30}}{12}.$$  

Analogously, a continuous time scale based on quarterly observed data is constructed via

$$\text{TIME}\cdot Q_i = \text{YEAR}_i + \frac{\text{QUARTER}_i - 1 + \frac{45 - 1}{90}}{4}$$

where QUARTER$_i$ is a variable running from 1 to 4 indicating the quarter of transaction. Panel (b) shows three indexes: One is based on daily observations (i.e., it uses the exact transaction dates), one is based on monthly and the last one on quarterly observations. Considering the indexes based on daily and monthly observations, there are hardly any deviations. In fact, the average absolute deviation is only 0.00165. The indexes based on daily and quarterly observations deviate much stronger, which is not surprising as the latter index can not detect changes within quarters. Despite that, the general development as well as turning points are identical. Given the fact that this index relies on very imprecise temporal information, its accuracy is astonishing.

The reliability of an index increases with rising sample sizes. Here, the number of observations is very large but what happens to the index if there were less observations? To answer this question, I create random sub-samples by sampling without replacement and compare indexes based on these sub-samples to the full-sample index. Deviations are very small: the average absolute deviation ranges between 1.91% in case of the 10% and 0.26% in case of the 90%-sub-sample. Panel (c) shows indexes relying on 10% and 25% of the observations. The indexes based on such small sample sizes (the 10%-sub-sample consists of less than 30,000 observations over a time span of eleven years!) still detect the same general patterns in the housing market in terms of overall level, major turning points, peaks and troughs. This findings suggest that the proposed methodology delivers reliable results also for strikingly small sample sizes.
6 Conclusions

This paper proposes an extension of the classical hedonic time-dummy method to construct house price indexes. Instead of period-wise indicators a smooth function controls for the temporal effect delivering an index that can be evaluated arbitrarily frequently. Such high-frequency indexes can serve as an early warning tool that detects movements in the housing market in “real-time” and measures changes precisely. The index at hand is also perfectly suited to act as an underlying for property derivatives. Though the residential property derivatives market is not strongly developed yet, there is potential for derivatives that allow investors to increase their exposure to the housing market without having to purchase a house.

The proposed continuous index outperforms discrete indexes in several ways: Discrete indexes are sensitive towards the choice of period lengths and starting points whereas continuous indexes are not making them more objective. In general, median and time-dummy indexes approach the continuously estimated index with decreasing period lengths. When periods are shortened too much, the indexes become very wiggly and unreliable tough. Next to these objectivity properties, the continuous index has some appealing robustness properties: Time-fixity holds, the index delivers reliable results also when the exact transaction date is not known but only the month or quarter of transaction and the methodology is applicable also in case of strikingly small sample sizes.

Next to a standard approach, a method allowing for flexible shadow prices is proposed. It turns out that for this dataset changes in the valuation of house characteristics hardly affect the index whereas there are significant changes in the locational effect over time.

Acknowledgements

This project has benefited from funding from the Austrian National Bank (Jubiläumsfondsprojekt 14947) and the JungforscherInnenfonds sponsored by the council of the University of Graz. I thank Australian Property Monitors for supplying the data used in this paper.

A preliminary version of this paper was presented at the inaugural conference of the Society of Economic Measurement at the University of Chicago Booth School of Business in 2014. I thank Jan de Haan (Delft University of Technology/Statistics Netherlands), Robert J. Hill (University of Graz), Robert Inklaar (University of Groningen), Alicia Rambaldi (University of Queensland), D.S. Prasada Rao (University of Queensland) and Michael Scholz (University of Graz) for their valuable and helpful comments.
References


Divisia, F. (1926), L’indice Monétaire et la Théorie de la Monnaie, Sirey, Paris.


Notes

[1] For reasons of convenience, 30 days per month are assumed. The benefits of using the true time scale or working days only are negligible.

[2] Ideally, the evaluation frequency should coincide with the format of the transaction date, e.g., if the exact transaction date is known a daily evaluation shall be preferred over a weekly evaluation.

[3] Using Simon Wood’s R-package mgcv such a matrix can be easily obtained by the function predict.gam(..., type='lpmatrix'). In a time-dummy setting the matrix $M$ is then given by

$$M = \begin{pmatrix}
1 & 0 & \ldots & 0 & - \bar{x}^T & - \\
0 & 1 & \ldots & 0 & - \bar{x}^T & - \\
\vdots & \vdots & \ddots & \vdots & \vdots & \vdots \\
0 & 0 & \ldots & 1 & - \bar{x}^T & - \\
\end{pmatrix}.$$

When smooth components enter the model, $M$’s structure is a bit more complicated but remains conceptually identical to the time-dummy case.


[5] Six months are chosen as minimum time span between two transactions following usual methodological choices for repeat-sales indexes (see for instance S&P/Case-Shiller Home Price Index: S&P Dow Jones Indices, 2015).

[6] State/Territory Land Titles Office, Valuers’-General Office or similar equivalent

[7] To guarantee comparability, all indexes are normalized with respect to their starting period / point.
Appendix

A Estimating AMs

This appendix summarizes the theory of Additive Models (AMs) and gives details regarding the choices made in this paper. Wood (2006) offers a generous documentation of AM theory and Wood’s R-package `mgcv` is a valuable tool to use these models in practise.

A.1 AMs in a nutshell

AMs are linear models that additionally include non-linear, non-parametrically estimated effects. These effects may be one- or multidimensional, e.g.,

\[ y_i \sim \alpha + \beta x_{1i} + f_1(x_{2i}) + f_3(x_{3i}, x_{4i}) \]

The functions \( f_j \) are a priori unknown and only ascertained in the estimation process. The functional form of the effects are determined by the data and not by the statistician; hence, such models increase objectivity.

To ease notation, let’s assume that the model consists of a single one-dimensional smooth function, i.e.,

\[ y_i \sim f(x_i). \]  \hspace{1cm} (5)

The function \( f \) is represented as a linear combination of given basis functions, i.e.,

\[ f(x) = \sum_{k=1}^{K} b_k(x) \beta_k. \]  \hspace{1cm} (6)

The basis functions \( b_k \) are known, but the parameters \( \beta_k \) are estimated. “Estimating the smooth function” therefore essentially means “estimating the parameters \( \beta_k \).” From this representation it is evident that model (5) can be written as linear model,

\[ y_i \sim f(x_i) = \sum_{k=1}^{K} b_k(x_i) \beta_k, \text{ and hence } y \sim X\beta, \]

where \( X \) is the design matrix with entries \( b_k(x_i) \) in the \( k \)th row and \( i \)th column. Usually, the chosen basis functions are splines. The choice of the type of splines used is a fundamental part of the model specification. In this paper, I use an adaptive basis approach based on P-splines for the one-dimensional time effect and thin-plate regression splines for the bivariate locational effect. The benefits of these choices are listed below.

A very convenient way to control the smoothness of the function \( f \) is to include a penalty that measures \( f \)’s wiggliness. Let \( J(f) \) be such a penalty function. Instead of fitting the model by just minimizing the sum of squared errors

\[ ||y - X\beta||^2 \]
a penalized version of the latter

\[ \|y - X\beta\|^2 + \lambda J(f) \]  

(7)
takes over. The smoothing parameter \( \lambda \) controls for the trade-off between model smoothness and model fit. \( \lambda \to \infty \) leads to a straight line estimate which is the smoothest possible outcome and \( \lambda = 0 \) to an unpenalized fit. Hence, \( f \)'s smoothness

\( f \) is a function of the single parameter \( \lambda \).

There are two important observations: First, \( f \) is linear in the coefficients \( \beta \) and, second, the most convenient penalty functions measure \( f \)'s wiggliness as a quadratic form in the coefficients. Hence, penalties can usually be written as

\[ J(f) = \beta^T S \beta, \]

where \( S \) is a matrix of known coefficients that characterize the penalty used. Altogether the minimization problem is given by

\[ \min_{\beta} \left( \|y - X\beta\|^2 + \lambda \beta^T S \beta \right). \]  

(8)

For a given \( \lambda \) the solution is

\[ \hat{\beta} = (X^T X + \lambda S)^{-1} X^T y. \]

The smoothing parameter \( \lambda \) enters the model as additional parameter. I use a selection criterion based on maximum marginal likelihood optimization (see Wood, 2011, for more details). This criterion is particularly preferable when using adaptive smoothers as I do in this paper.

A.2 Adaptive smoothers

Eilers and Marx (1996) suggested combining a basis consisting of B-splines (De Boor, 1978) and a penalty based on second order differences which is known as P-splines. Each B-spline of degree \( q \) is positive on a domain spanned by \( q + 2 \) knots and zero otherwise. Hence, B-splines are local basis functions, i.e., B-splines are able to model local fluctuations accurately. Let \( f \) be represented as a linear combination of B-splines \( b_k \) and coefficients \( \beta_k \) as in (6). Eilers and Marx (1996) measure \( f \)'s wiggliness by the sum of squared second differences of the coefficients, i.e.,

\[ J(f) = \sum_{k=2}^{K-1} (\beta_{k-1} - 2\beta_k + \beta_{k+1})^2. \]

They argue that this penalty is a good discrete approximation to the integrated square of the \( q \)th derivative, which is the standard penalty function. Wood (2011) names “the ease with which they [P-splines] can be modified to perform non-standard smoothing tasks [such as adaptive smoothing], at relatively little loss of performance relative to more computationally complex smoothers” as a major advantage of P-splines.
In the case of adaptive smoothers, the strength of the penalty varies with the covariate, i.e., there is not a single global smoothing parameter but rather several ones depending on the value of the covariate. This flexibility is particularly useful when the speed at which \( f \) changes varies in different regions of the covariate’s domain. Therefore, weights are included into the penalty function:

\[
J^{ad}(f) = \sum_{k=2}^{K-1} w_k (\beta_{k-1} - 2\beta_k + \beta_{k+1})^2.
\]

The weights \( w_k \) depend on the index \( k \) that numbers the basis functions and hence on the covariate itself. The unknown weights are modelled as a smooth function of \( k \) using another B-spline basis. The resulting penalty can be written as a sum of penalties multiplied by smoothing parameters. The optimization problem is formulated as in (8) and solved as before. See Wood (2011) for more details and further references.

Housing markets experience both, times of relatively stable prices and times of rapid changes. In other words, the speed of change in prices is not constant over time. A single smoothing parameter is therefore expected to be too restrictive. The time effect was hence estimated using an adaptive P-splines approach. The dimension of the smoothing basis \( K \) was chosen to be 60, whereas the dimension of the penalty basis was chosen to be 6. That means that six smoothing parameters enter the smooth, which estimates range between 0.002 and 1.454.

### A.3 Thin-plate regression splines

Thin-plate regression splines (TPRS) are a convenient class of smoothers to model multivariate covariates. These smoothers are based on thin-plate splines (Duchon, 1977) that use penalties of the form

\[
J(f) = \int \int \left( \frac{\partial^2 f}{\partial x_1^2} \right)^2 + 2 \left( \frac{\partial^2 f}{\partial x_1 \partial x_2} \right)^2 + \left( \frac{\partial^2 f}{\partial x_2^2} \right)^2 \, dx_1dx_2.
\]

The optimization problem (7) together with this penalty has a closed form solution, which can be used to write (7) again as a sum of an unconstrained sum of squares and a quadratic form multiplied by a smoothing parameter. Hence, the methods used to estimate a penalized regression model can be applied.

The main advantage of thin-plate splines is that neither basis functions nor the position of knots have to be specified. Both emerge naturally as a consequence of the choice of penalty. Thin-plate splines are suited to model multidimensional predictor variables. However, these advantages come at the cost of a high level of computational resource consumption. Wood (2006) overcomes this problem by introducing a low-rank approximation of thin-plate splines and calls the result TPRS. This approximation retains the most important advantages of thin-plate splines and reduces computational cost guaranteeing practicability.

I use TPRS to estimate the bivariate locational effect. This choice is particularly beneficial as TPRS are rotational invariant. This isotropy property is desirable.
when modelling geographical coordinates. Moreover, TPRS are usually the best choice when all predictor variables are measured in the same unit as coordinates are (i.e., in degrees). In the hedonic equation I used a basis dimensionality of 600. A sensitivity analysis suggested that this choice is appropriate and further increasing the dimension did not have noticeable effects on the resulting price index.

B Complete-case analysis

[Figure 7 about here.]

In this paper I perform a complete case analysis, i.e., I exclusively rely on the fully recorded observations. This is appropriate as the estimated house price distributions using the complete and full dataset respectively are almost identical after controlling for location and time.\cite{8} This is seen in Figure 7: The figure shows estimated densities of house prices separately for regions\cite{9} and years based on the full and the complete datasets. Eight out of 88 possible graphs are depicted. Consistently, both densities are almost identical. This is also true for the huge majority of the other 80 possible combinations indicating that – after controlling for locational and temporal effects – the price distribution does not depend on the completeness of observations.
Figure 1: Left: Estimated locational effect. Right: Continuous index with point-wise standard error bands ($\pm 2 \cdot \hat{s.e.}$).
Figure 2: Shortcomings of discrete indexes.
Figure 3: Continuous index compared to (a) time-dummy and median indexes and (b) time-dummy indexes based on very short period lengths.
Figure 4: (a) Daily and yearly evaluated continuous index, (b) Stratification index compared to continuous index.
Figure 5: Flexible shadow prices. Top: Comparing indexes with different degrees of flexibility. Bottom: Updating structure for $f(\text{LONG}, \text{LAT})$. 

<table>
<thead>
<tr>
<th>Year</th>
<th>2001</th>
<th>2002</th>
<th>2003</th>
<th>2004</th>
<th>2005</th>
<th>2006</th>
<th>2007</th>
<th>2008</th>
<th>2009</th>
<th>2010</th>
<th>2011</th>
</tr>
</thead>
<tbody>
<tr>
<td>Model 1</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
<td>5</td>
<td>6</td>
<td>7</td>
<td>8</td>
<td>9</td>
<td>10</td>
<td>11</td>
</tr>
<tr>
<td>Model 2</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
<td>5</td>
<td>6</td>
<td>7</td>
<td>8</td>
<td>9</td>
<td>10</td>
<td>11</td>
</tr>
<tr>
<td>Model 3</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
<td>5</td>
<td>6</td>
<td>7</td>
<td>8</td>
<td>9</td>
<td>10</td>
<td>11</td>
</tr>
<tr>
<td>Model 4</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
<td>5</td>
<td>6</td>
<td>7</td>
<td>8</td>
<td>9</td>
<td>10</td>
<td>11</td>
</tr>
<tr>
<td>Model 5</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
<td>5</td>
<td>6</td>
<td>7</td>
<td>8</td>
<td>9</td>
<td>10</td>
<td>11</td>
</tr>
<tr>
<td>Model 6</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
<td>5</td>
<td>6</td>
<td>7</td>
<td>8</td>
<td>9</td>
<td>10</td>
<td>11</td>
</tr>
</tbody>
</table>
Figure 6: Robustness analysis: (a) time fixity, (b) frequency, (c) sample sizes.
Figure 7: House price densities: full vs. complete dataset.

Miriam Steurer, Caroline Bayr: Quantifying Urban Sprawl using Land Use Data

Porfirio Guevara, Robert J. Hill, Michael Scholz: Hedonic Indexes for Public and Private Housing in Costa Rica

Sofie R. Waltl: A Hedonic House Price Index in Continuous Time

Ronald Wendner: Do Positional Preferences for Wealth and Consumption Cause Inter-Temporal Distortions?

Maximilian Gödl, Christoph Zwick: Stochastic Stability of Public Debt: The Case of Austria

Gabriel Bachner: Land Transport Systems under Climate Change: A Macroeconomic Assessment of Adaptation Measures for the Case of Austria

Thomas Schinko, Judith Köberl, Birgit Bednar-Friedl, Franz Pretenthaler, Christoph Töglhofer, Georg Heinrich, Andreas Gobiet: An Integrated Economic Analysis of Climate Change Impacts on Tourism in Austria Across Tourism Types and Seasons

Sugata Ghosh, Ronald Wendner: Positional Preferences, Endogenous Growth, and Optimal Income- and Consumption Taxation

Stan Metcalfe: George Shackle and The Schumpeterian Legacy.

Birgit Bednar-Friedl, Karl Farmer: Existence and efficiency of stationary states in a renewable resource based OLG model with different harvest costs

Karl Farmer, Irina Ban: Modeling financial integration, intra-EMU and Asian-US external imbalances

Y. Hossein Farzin, Ronald Wendner: The Time Path of the Saving Rate: Hyperbolic Discounting and Short-Term Planning

Robert J. Hill, Iqbal A. Syed: Hedonic Price-Rent Ratios, User Cost, and Departures from Equilibrium in the Housing Market

Christian Gehrke: Ricardo’s Discovery of Comparative Advantage Revisited

Sabine Herrmann, Jörn Kleinert: Lucas Paradox and Allocation Puzzle – Is the euro area different?

Christoph Zwick: Current Account Adjustment in the Euro-Zone: Lessons from a Flexible-Price-Model

Karl Farmer: Financial Integration and EMUs External Imbalances in a Two-Country OLG Model

Caroline Bayr, Miriam Steurer, Rose-Gerd Koboltschnig: Scenario Planning for Cities using Cellular Automata Models: A Case Study

Y. Hossein Farzin, Ronald Wendner: Saving Rate Dynamics in the Neoclassical Growth Model – Hyperbolic Discounting and Observational Equivalence

Maximilian Gödl, Jörn Kleinert: Interest rate spreads in the Euro area: fundamentals or sentiments?

Christian Lininger: Consumption-Based Approaches in International Climate Policy: An Analytical Evaluation of the Implications for Cost-Effectiveness, Carbon Leakage, and the International Income Distribution

Veronika Kulmer: Promoting alternative, environmentally friendly passenger transport technologies: Directed technological change in a bottom-up/top-down CGE model

Paul Eckerstorfer, Ronald Wendner: Asymmetric and Non-atmospheric Consumption Externalities, and Efficient Consumption Taxation

Michael Scholz, Stefan Sperlich, Jens Perch Nielsen: Nonparametric prediction of stock returns with generated bond yields
09–2012  Jörn Kleinert, Nico Zorell: The export-magnification effect of offshoring

08–2012  Robert J. Hill, Iqbal A. Syed: Hedonic Price-Rent Ratios, User Cost, and Departures from Equilibrium in the Housing Market

07–2012  Robert J. Hill, Iqbal A. Syed: Accounting for Unrepresentative Products and Urban-Rural Price Differences in International Comparisons of Real Income: An Application to the Asia-Pacific Region


05–2012  Miriam Steurer, Robert J. Hill, Markus Zahrhofer, Christian Hartmann: Modelling the Emergence of New Technologies using S-Curve Diffusion Models

04–2012  Christian Groth, Ronald Wendner: Embodied learning by investing and speed of convergence

03–2012  Bettina Brüggemann, Jörn Kleinert, Esteban Prieto: The Ideal Loan and the Patterns of Cross-Border Bank Lending

02–2012  Michael Scholz, Jens Perch Nielsen, Stefan Sperlich: Nonparametric prediction of stock returns guided by prior knowledge

01–2012  Ronald Wendner: Ramsey, Pigou, heterogenous agents, and non-atmospheric consumption externalities