Residential Property Price Indices for Austria[1]

Wolfgang Brunauer
Real(e)value ImmobilienbewertungsGmbH
Wolfgang Feilmayr
Vienna University of Technology

Abstract

This paper describes two residential property price indices (RPPI) available in Austria. The first one is a classical hedonic property price index with chained dummy evaluation, available for Vienna since 1986. The second one is what we call a "spatial imputation index". It has the advantage that is evaluable on arbitrary spatial scales above census tract level, and can take into account different perspectives for aggregation.

Challenges and Methods for Index Construction

Real estate consists of a wide range of characteristics that make each object unique. Therefore, households implicitly select a set of values for each of the K (structural and locational) characteristics $z = (z1, z2, \ldots, zK)$. The price of the house or apartment is a function of the whole bundle of characteristics, or the hedonic price function $P = P(z)$, with functional relationships possibly varying over time and space. This implies the following challenges that have to be taken into account during the construction of a RPPI:

- Comparability: The index must take into account different quality levels (e.g. location, size, age, technical equipment) and make the results comparable across time periods.
- Change of standards: To construct a representative index, it is necessary to consider changes in quality standards over time and to rely on comparable characteristics. If price indices are evaluated for nonrepresentative characteristics, this may result in biased results.

[1] Large parts of this paper are taken from Brunauer, Feilmayr and Wagner (2012).

- Time-varying effects: Even if the effects of price-determining characteristics change only gradually, in the long run they cannot be assumed to be fixed.
- Representativeness: During construction and aggregation of sub-indices, it is necessary to take into account the underlying population we want to describe.

The main methods to control for changes in property characteristics discussed in the literature are *stratification, repeat sales methods* and *hedonic regression methods* (for more details, see European Commission 2011, chapters 3 and 4).

In the following, we concentrate on hedonic regression models. In the literature, two subgroups of hedonic indices are discussed. The time dummy variable method, which models the property's price as a function of its characteristics and a set of time dummy variables, is most popular. The first index discussed in this paper is such a time dummy index. However, as data on all sample periods are pooled, the resulting indices may be subject to revisions. Furthermore, the results implicitly correspond to a transaction-weighting. To deal with these problems, hedonic imputation methods were developed. The second method presented here is a special type of such an imputation index.

The Austrian RPPI

Both RPPI presented in this paper rely on quotation prices, which include price observations combined with "structural" (property-specific) variables. Approximately 10,000 observations are available each year from the internet platform AMETA-NET (EDI-ORG, Linz). "Locational" (location-specific) variables are used to adjust for spatial differences down to census tract level[2]. On the basis of the available data, indices are estimated for apartments and single family houses (SFH).

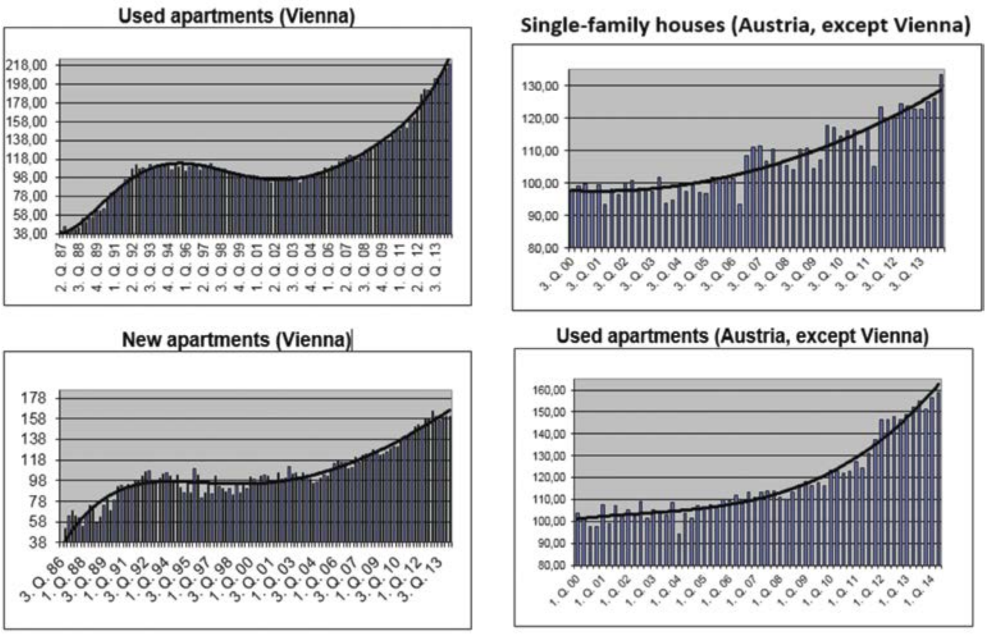
Time Dummy Index

The time dummy index applies multiple linear regression models, where (log) price is explained by building attributes and district dummies. Overlapping models are estimated over a period of one year, with dummy-coded quarters, where the resulting time effects are chained over the last quarter of the previous model. In total, six separate models are estimated, split by object type and region:

- "Vienna" and "Austria, except Vienna"
- SFH, apartments (new), apartments (used)
- The total effect is determined by a weighting scheme (see below)

[2] In Austria, there are 8,748 census tracts in 2,379 municipalities, which form 121 districts across the 9 Austrian provinces.

Chart 1: Selected Results for Sub-Models of the Time Dummy Index[1]



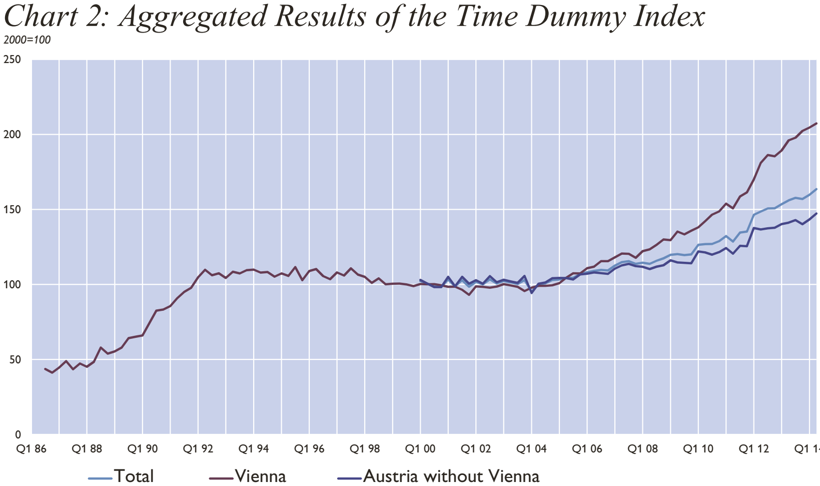
Source: Authors' calculations.

[1] *Results available from http://www.srf.tuwien.ac.at/feil/immobilienbewertung/Index214.pdf.*

The sub-indices are weighted by numbers of transactions from the land register (2008-2013).

Chart 2 shows the weighted results for "Vienna" and "Austria without Vienna" and the overall index development ("Total").

Chart 2: Aggregated Results of the Time Dummy Index



Source: OeNB, Feilmayr, Department of Spatial Planning, Vienna University of Technology.

However, there are some drawbacks of this method: Due to model construction, the time dummy index faces the problem of "retrospective" variation of time effects. Furthermore, changes in quality levels are not explicitly considered in the model. And finally, as described above, the model implicitly incorporates a transaction perspective (or more precisely, a weighting by the number of quotation prices). Therefore, there is no inference for unobserved regions.

Spatial Imputation Index

The second index described in this paper, the spatial imputation index, tries to overcome the shortcomings of the time dummy index. The basic principle of this index is to estimate hedonic models for rolling windows of two years, evaluating the results of these models over all census tracts in Austria for representative bundles of characteristics, and comparing the results of the current quarter to a base quarter.

Another issue in hedonic price modeling is the occurrence of nonlinear functional relationships and unexplained spatial heterogeneity, i.e. spatially varying relationships that cannot be explained by location-related variables. The time dummy index is based on classical linear models, which allow for nonlinearity in covariate effects only in a very restricted fashion (by integrating polynomial terms in the equation), again possibly resulting in functional misspecification. Furthermore, unexplained spatial variation is modeled by dummy effects. This approach yields very unstable results in districts where only few observations are recorded. Also, the time index is modeled according to the dummy approach, which results in very volatile effects if, in a certain time period, there are relatively few or outlying observations.

As linear models may seriously impair index construction, the new model uses a semiparametric approach, namely Generalized Additive Models (GAMs) as described in Wood (2006). GAMs model continuous covariates using penalized regression splines, which allow for nonlinearity in a regularized statistical framework. Distributional and structural assumptions, given covariates and parameters are based on Generalized Linear Models (GLMs), $E(y_i|z_i,x_i)=h(\eta_i)$[3]. However, instead of a linear predictor, GAMs apply the additive predictor $\eta_i=f_1(z_{i1})+\ldots+f_q(z_{iq})+x_i'\gamma$, where $x'\gamma$ is the usual parametric part of the predictor, z_j is a continuous covariate, time scale or district index and f_j are (not necessarily continuous) functions of these covariates. The trade-off between data fidelity and smoothness is governed by model selection criteria, in our case the generalized cross-validation criterion (Wood, 2006). Unexplained spatial heterogeneity is modeled by random effects. Random effects penalize the lack of information in a district: The fewer obser-

[3] For the models applied here, we use the identity link, i.e. the linear predictor corresponds to the mean of the distribution function.

vations there are in one unit, the more they tend toward the "baseline" effect of the model, i.e. the level predicted by location-specific levels of covariates. Again, penalization is achieved by model selection criteria.

For the current version of the RPPI, the log of the property price is taken as the dependent variable. The effect of the age of the respective object, the specific living area and, in the case of family homes, also the plot area as well as locational covariates[4] are modeled as penalized regression splines, and unexplained spatial heterogeneity is modeled by district random effects[5]. The time trend is considered in five different ways:

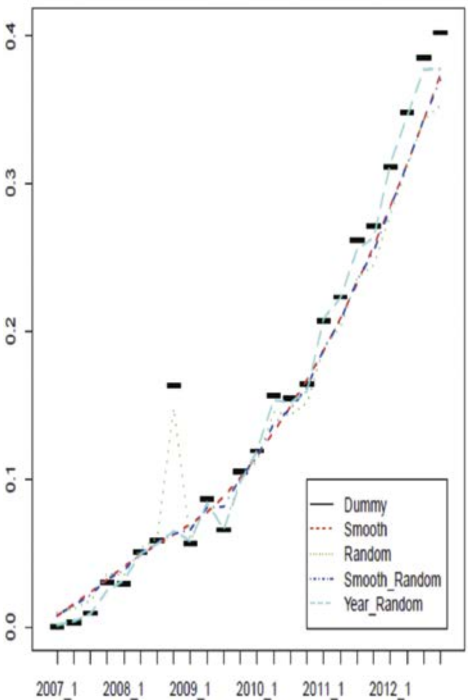
1. In the *dummy approach,* quarterly time dummy effects are estimated like in the old model. However, as can be seen in charts 1 and 2, this leads to rather volatile estimation results, e.g. for condominiums in Q4 08 and for family homes in Q4 06.
2. The second version is a *smooth* trend estimation, where a nonlinear time trend is estimated using penalized regression splines. However, for both models, this approach seems to underestimate abrupt jumps.
3. The third version estimates a *random* time effect. As mentioned before, random effects models tend to weight the estimated parameters toward zero. Although this helps prevent large outliers, the effect is underestimated if the market truly tends up- or downward over time.
4. The fourth model combines a nonlinear time trend and random effects for modeling deviations from this trend, which is why it is called *"smooth-random"* model. This approach should account for abrupt deviations from a continuous baseline trend. Nevertheless, the results seem to be dominated by the nonlinear trend.
5. Finally, a model is estimated that integrates time effects in a hierarchical manner: The baseline time effect is estimated as a yearly effect, and deviations from that effect are captured by random effects on a quarterly basis, which is why this model is called *"year-random"* model. It seems to provide a robust trade-off between data fidelity and penalization, as the baseline effect is defined by a relatively large subsample while abrupt changes are also captured to some extent.

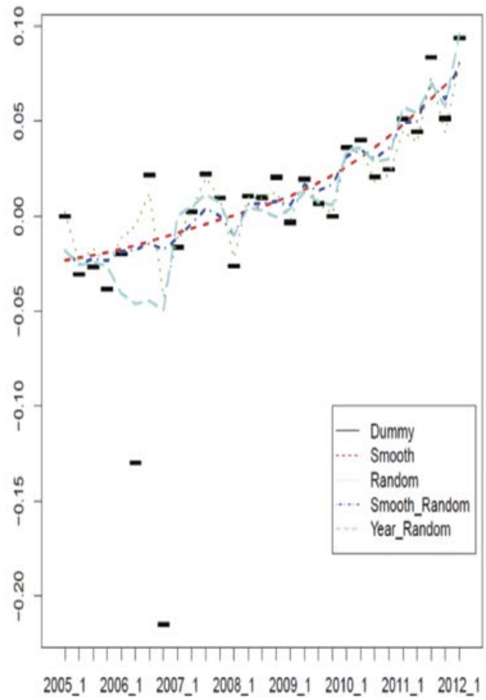
[4] For the new RPPI, the log of the share of academics or high school graduates in the total population of a census tract is taken as a proxy for the social composition of the neighborhood. Furthermore, four variables enter the model at the municipality level: an index for the price of building land; the number of overnight stays (as a proxy for touristic activity in the municipality); a measure for accessibility (as a measure of centrality); and the growth of purchasing power from 2005 to 2009.

[5] There are 121 census districts in Austria.

A comparison of these approaches is shown in chart 3. The subsequent chart 4 shows some nonlinear effects in the base model, where the central line is the estimated effect, and the grey areas correspond to 95% confidence intervals. In both cases, the explanatory variable is printed on the x-axis, while the y-axis can be approximately interpreted as the percentage effect on market values.

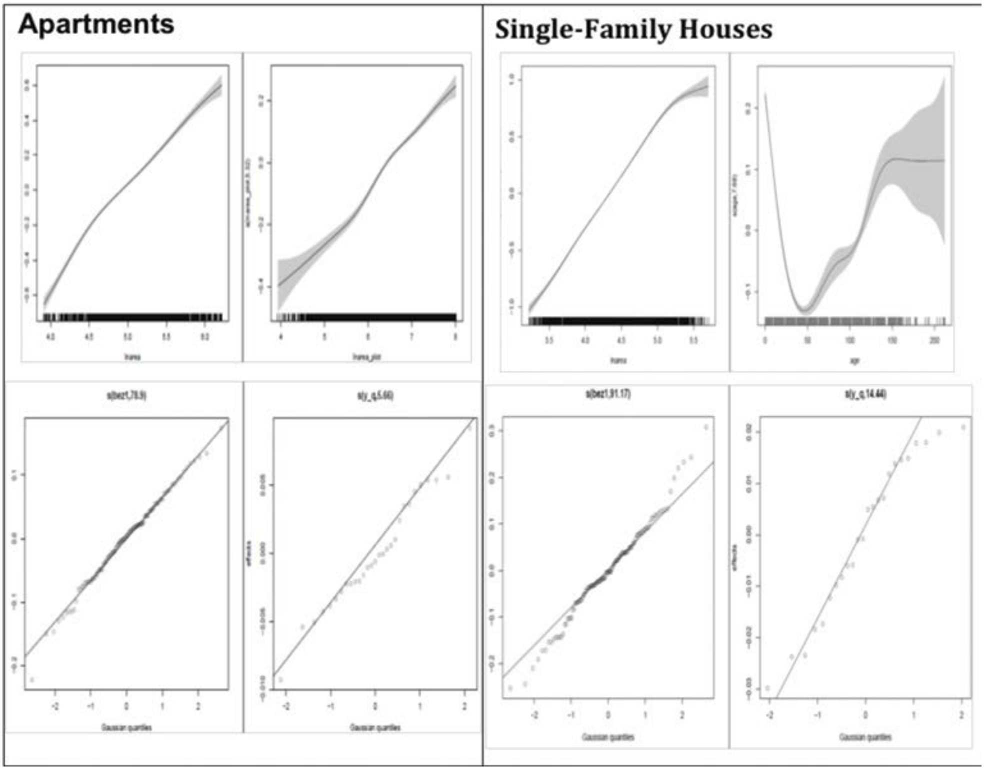
Chart 3: Various Trend Models for Single Family Houses and Apartments





Source: Authors' calculations.

Chart 4: Nonlinear Effects



Note: In clockwise order for SFH (left panels): Effect of log(floor area), log(plot area), random effects for districts, random effects for quarters. Apartments (right): Effect of log(floor area), building age, random effects for districts, random effects for quarters.

Using this methodological background, the following steps are taken to construct the spatial imputation index:

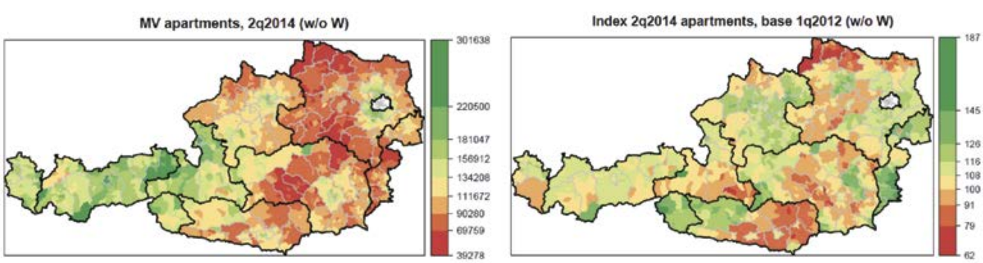
1. Definition of subsamples for model construction
   a) Base model: 2007Q1–2012Q1 (apartments) and 2005Q1–2012Q1 (SFH), with base period 2012Q1
   b) Comparison model: Moving window of two years, with comparison the latest available quarter
2. Determining "average characteristics" for the last year until the base and comparison period, respectively
3. Regression model for base and comparison period
4. "Spatial imputation":
   a) Simulation of an „average characteristics" object for each census tract in base and comparison period

   b) Leveraging on the similarities modeled by location covariates and the hierarchical structure for spatial heterogeneity
5. Index calculation on census tract level:
   a) Laspeyres-Index: Determine average characteristics of base period, impute these effects on census tract level
   b) Paasche-Index: Determine average characteristics of current period, impute these effects on census tract level
   c) Fisher-Index: Geometric mean of Laspeyres and Paasche
6. Rescaling to 2007Q1 = 100 using dummy effects from base model
7. Weighting scheme:
   a) For each submodel and state: Averaging over census tracts (except: SFH/ Vienna, "inner" districts)
   b) Average over state/type index values weighted by share of state/type on total stock (based on Household Finance and Consumption Survey/HFCS 2010)

The approach described above has the advantages that space is taken into account by location covariates on census tract and municipality level ("observed" spatial heterogeneity) and a hierarchical structure of state and district ("unobserved" spatial heterogeneity. Furthermore, representativeness is obtained on the one hand by imputation in each census tract, with spatially varying census tract characteristics, and a weighted aggregation using shares of respective state in total building stock.

In the following charts, we evaluate the results on census tract level with base period 1Q2012[6]. For both apartments and SFH, the spatial indices are strongly heterogeneous.

Chart 5: Results for Apartments on Census Tract Level for Austria Except Vienna

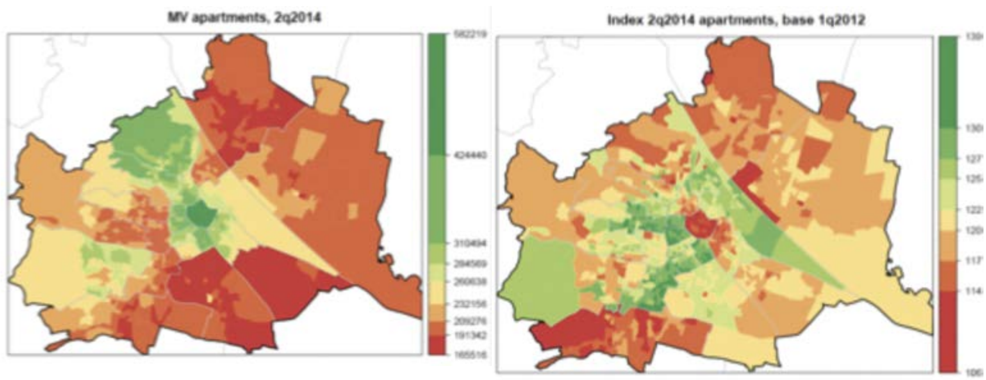


Note: Upper left panel: market value in base period, lower panel: Fisher index, upper right panel: market value in comparison period, lower panel Fisher index.

[6] Note that for the final index, the results are re-scaled to 1Q2007 as described above.

For apartments in Austria except Vienna, positive developments can be seen predominantly in urban and suburban regions as well as in the western parts of Austria with distinct tourism.

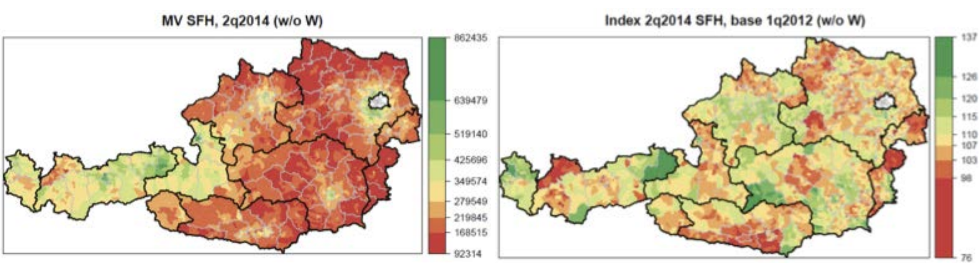
Chart 6: Results for Apartments on Census Tract Level for Vienna



Note: Left panel: market value in current (comparison) period; right panel: Fisher index.

For apartments in Vienna, we do not find any negative developments. The strongest positive developments can be found in the CBD and on the fringe of Vienna.

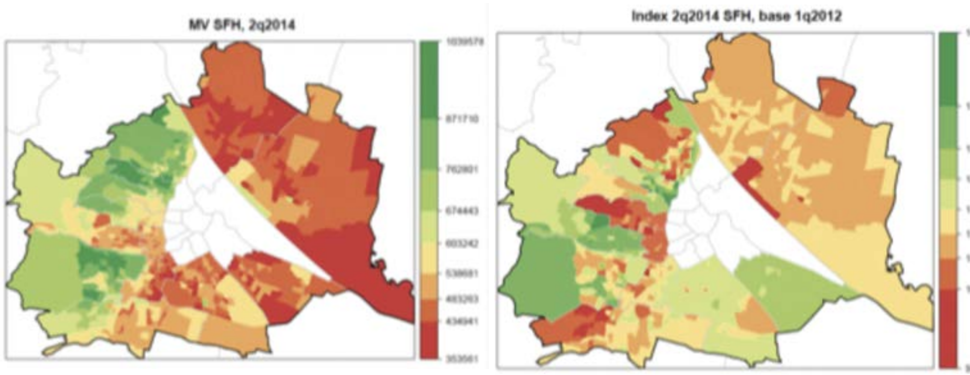
Chart 7: Results for SFH on Census Tract Level for Austria except Vienna



Note: Left panel: market value in current (comparison) period; right panel: Fisher index.

For SFH in Austria except Vienna, we find the strongest positive developments in the eastern part of Tyrol and the northern part of Vorarlberg. Also upper Austria and the very west of Styria show strong price increases. However, the latter is possibly due to statistical artifacts and still has to be examined in more detail.

Chart 8: Results for Apartments on Census Tract Level for Vienna (Except Inner Districts)



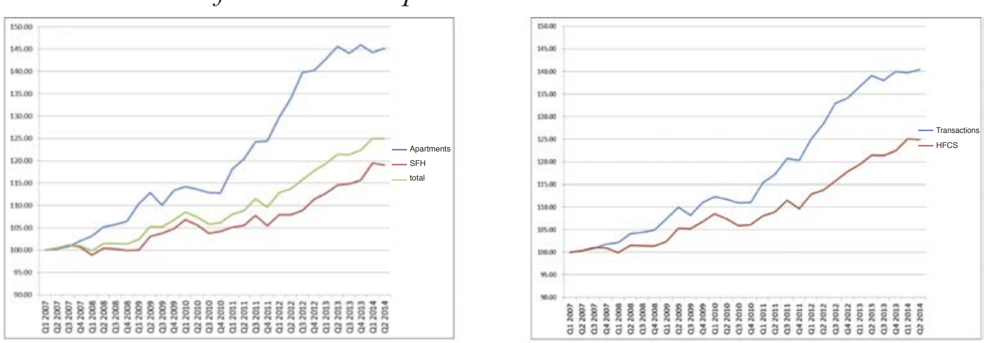
Note: Left panel: market value in current (comparison) period; right panel: Fisher index.

For SFH in Vienna, we exclude the inner districts from index construction as there is no market for SFH. Due to very small sample sizes the results for the single districts should be interpreted very carefully.

In a next step, the resulting index values on census tract level are aggregated on state level. The indices on state level are then weighted according to their share on total stock based on Household Finance and Consumption Survey/HFCS 2010.

The left panel of chart 9 shows the resulting index for apartments, SFH and the overall index. However, it is important to be aware of the effect of the weighting scheme: the right panel shows a comparison between the index resulting from the stock weighting scheme and an index resulting from transaction weights.

Chart 9: Results for SFH and Apartments



Note: Left panel: Index development for apartments, SFH and the overall index based on stock weighting scheme (based on HFCS 2010); right panel: comparison between stock and transaction weights.

Conclusion

In this paper, we describe two indices available in Austria: The time dummy index, which is available for Vienna for the period since 1986 and for Austria except Vienna since 2000, and the spatial imputation index, available for the period since 2007. Both indices are updated quarterly and available from the OeNB's website.[7] While the spatial imputation index solves some problems that are obvious for the time dummy index, there is still much room for improvement concerning data, model and weighting scheme:

1. Data: The spatial imputation model is based on quotation/asking price data. Obtaining data sources that are not affected by weaknesses of asking price data would be preferable.
2. Model: The underlying model could be improved e.g. by spatially varying effects for object characteristics and by including location covariates measured in respective year (currently, locational covariates only provide a "cross-section" perspective, i.e. they do not vary over time). Additional covariates could be included measuring fundamental factors for demand (e.g. demographic development) and supply (e.g. available floor space).
3. Weighting scheme: Although the disaggregated index values overall give a plausible picture, there are some regions where statistical artifacts may occur. Furthermore, a more detailed weighting scheme seems appropriate, accounting for share of stock e.g. on district level.

Literature

Brunauer, W., Feilmayr, W., Wagner, K. 2012. A New Residential Property Price Index for Austria, Statistiken – Daten und Analysen – Q3/12. OeNB.

European Commission. 2011. Handbook on Residential Property Prices Indices. Retrieved from http://epp.eurostat.ec.europa.eu/portal/page/portal/hicp/methodology/owner_occupied_housing_hpi/rppi_handbook on July,19th 2012.

Wood, S. 2006. An Introduction to Generalized Additive Models with R. Boca Raton: Chapman and Hall.

[7] www.oenb.at/isaweb/report.do?lang=DE&report=6.6.