17th Ottawa group meeting
Rome, 7 – 10 June 2022

Bringing the Hedonic-repricing method up to date to adjust for qualitative differences in the residential real estate price index

Authors: Corinne Becker Vermeulen, Manuel Brand, Swiss Federal Statistical Office

Abstract

The hedonic repricing method is not always considered an appropriate method for adjusting the quality of real estate transactions to measure only pure price movements, is not explicitly mentioned in all international recommendation manuals and is even sometimes criticised, although it is still used by many countries for the production of the residential property price index (RPPI). What is this method really about and how can it be improved? This paper analyses the quality adjustment method chosen by Switzerland for the quarterly production of the house price index, compares it with other methods and tries to reduce the main weakness of this method.

1. The Swiss residential property price index in brief[1].

Switzerland was one of the last European countries to produce an official house price index. Why is this? Switzerland is primarily a country of tenants[2], as home ownership is extremely expensive. The development of property prices has also been calculated for decades by private and banking institutes. Private supply in this very lucrative sector was abundant, however provided a mixed picture, with each institute having its own data source[3] and calculation method[4]. The Swiss Parliament, supported by the Swiss National Bank, the State Secretariat for Economic Affairs and the Financial Market Supervisory Authority, mandated the Federal Statistical Office to produce an official house price index, complying with international standards and good practices and with the quality of an official statistics.

In November 2020, the FSO published its residential property price index for the first time, with results backdated to 2019. The development of this new indicator was not without its challenges in terms of data sourcing. The 25 largest mortgage institutions (in fact the 25 largest Swiss banks) were identified as the best source of data, however banking secrecy placed many obstacles in our way, despite the legal basis of federal statistics and the law on statistical confidentiality.

[1] A full methodological description is available on the internet [Swiss residential property price index- Methodological principles.]

[2] 63.8% of tenants or cooperative members and 36.2% of owners according to FSO's occupancy status statistics (https://www.bfs.admin.ch/bfs/en/home/statistics/construction-housing/dwellings/housing-conditions/tenants-owners.html).

[3] These private institutes provide hedonic appraisal tools and use the transaction data that the banks evaluate herewith for the index calculation.

[4] Given the small amount of quarterly transactions, private institutes create a pool of data in order to be able to apply the characteristics prices method.

The Swiss residential property price index (IMPI) is calculated on the basis of approximately 7,000 real estate transactions per quarter. Given that mortgage institutions finances 90% of the transactions and that the 25 largest mortgage institutions in our sample represent almost 90% of the market if we consider the entire mortgage market, the coverage of our sample in relation to the total population is excellent. The variables used to compile the property price index are the following:

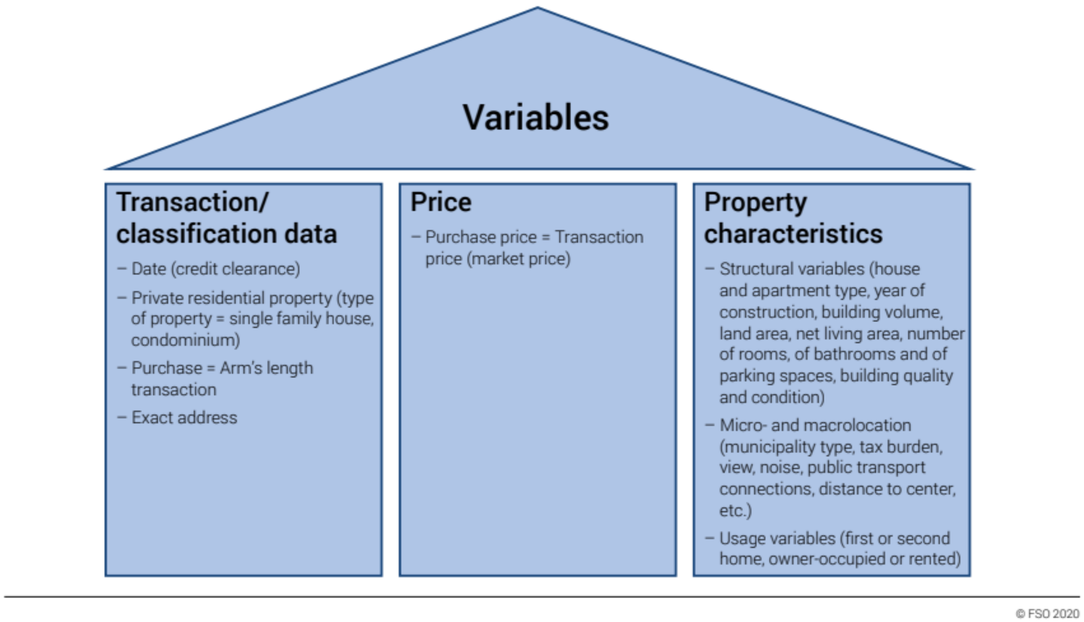

Figure 1: the data for the Swiss residential property price index

The variables relating to the transaction, the price, the structural and usage variables are supplied by the banks; the location variables (micro and macro location) come from external, publicly accessible geolocalised datasets[5] and are matched with the variables supplied by the data providers[6]. About twenty variables are thus available.

We have stratified our sample according to the type of property (single-family houses and condominiums) and according to five types of municipality (three types of urban municipalities, intermediate municipalities and rural municipalities), in order to reduce the heterogeneity of our sample and to be able to publish sub-indices. However, this stratification does not allow us to take into account all the variables influencing prices; this is why the stratification is combined with a hedonic model, which we will discuss in the next chapter.

[5] Publicly-accessible databases with geolocation data: – Train noise, road traffic noise (Federal Office for the Environment) – Travel times to centers, public transport quality, proportion of second homes per municipality (Federal Office for Spatial Development) – Aircraft noise (Federal Office of Civil Aviation) – Elevation model, lakes, rivers, high-voltage power lines, slope, exposure (Federal Office of Topography) – Tax burden (Federal Tax Administration) – Municipality type, major statistical regions (Federal Statistical Office) – Mountain and lake view (FSO's own view model derived from elevation models).

[6] The matching is carried out by the data provider itself, using specially developed software, on the basis of the address of the real estate object and the addresses of the register of buildings and dwellings. Once the matching has been completed, the address is deleted and the complete data is delivered to the FSO.

2. The choice of the hedonic model

Hedonic adjustment is the most frequently used quality adjustment method in the production of a house price index. Among the methods proposed and described in international handbooks[7], there are the time dummy hedonic method (rolling), the hedonic imputations method and the characteristics prices method. The hedonic repricing method is not always explicitly mentioned in the handbooks, as it is considered a variant of the imputation method.

The hedonic method consists in breaking down the total price of a property according to its characteristics and estimating a value for each of its characteristics (shadow price). We can thus estimate the price of an extra room, an extra square meter, the price for an open view and the price for a quiet environment. The set of characteristics corresponds to our basket and the shadow prices of the characteristics to the prices that we have to follow over time.

In choosing a model type, several factors must be taken into consideration: the number of transactions available in each period depending on the frequency of publication (quarterly if results are published quarterly or monthly if results are published monthly), the availability of structural, usage and possibly macro and micro location variables, the stability or instability of the national real estate market, as well as the willingness to publish sub-indices and to break down the total results into sub-indices.

The (rolling) time dummy hedonic method is relatively simple to apply. This approach is appropriate when there are few transactions. Data pooling from several periods allows to obtain coefficients/shadow prices for each characteristic of better quality than if they were calculated with data from a single period. However, the pooling comes with a price. The longer the window, the smoother will be the series[8]. The index will become less responsive and therefore loses timeliness. If we stratify the sample, a model must be estimated for each stratum and each period. Its (log-linear) formula is [9]:

$$\ln p_n^t = \beta_0 + \sum_{t=1}^{T} \delta^t D^t + \sum_{k=1}^{K} \beta_k Z_{nk}^t + \varepsilon_n^t$$

t - *period*
n - *dwelling n*
k - *characteristics*
$\ln p_n^t$ - *price logarithm*
β_0 - *intercept*
δ^t - *coefficient of the time dummy variable that will create the index*
D^t - *time dummy variables*
β_k - *shadow price of characteristic k*
Z_{nk}^t - *quantity of characteristic k in period t and dwelling n*
ε_n^t - *error term*

[7] Handbook on Residential Property Prices Indices (RPPIs), Eurostat 2013 KS-RA-12-022-EN.PDF (europa.eu) and Residential property price index (RPPI) practical compilation guide, IMF 2020 Residential Property Price Index (RPPI) Practical Compilation Guide (imf.org)
[8] Silver Mick (2018); How to measure hedonic property price indexes better.
[9] Residential property price index (RPPI) practical compilation guide, IMF 2020, page 39

The index for current period (t) is derived as follows:

$$I_t = \exp(\hat{\delta}_t) * 100$$

The characteristics prices method is appropriate when there are a large number of transactions in each period, because the value of the characteristics (shadow price) is estimated each quarter[10]. This method consists in defining an average basket of characteristics for the reference period (a typical property) and following the development of the prices of this basket over time by applying the values of each characteristic estimated at each period by the hedonic model. If we have geolocalised variables (micro or macro location), the use of this method becomes difficult because while it is easy to average structural variables, it is less reasonable to average geolocalised values. Since a regression is estimated each quarter, this method has the advantage of providing predicted prices that are up to date and thus very responsive to price changes. On the other hand, how to determine the reference basket is crucial, and if the average basket of characteristics is not updated regularly, it may no longer be representative of the real estate market, especially when it is dynamic. Again, if we stratify the sample, a model must be estimated for each stratum and each period. Its arithmetic formula is[11]:

$$P_{CP} = \frac{\hat{p}^t}{\hat{p}^0} = \frac{\sum_{k=0}^{K} \hat{\beta}_k^t (z_k^*)}{\sum_{k=0}^{K} \hat{\beta}_k^0 (z_k^*)}$$

$\hat{p}^t$ - *predicted price in period t*
$\hat{p}^0$ - *predicted price in period 0*
$\hat{\beta}_k^t$ - *shadow price of characteristic k in period t*
$\hat{\beta}_k^0$ - *shadow price of characteristic k in period 0*
z_k^* - *standardized characteristic k*

The hedonic imputation method (single or double imputation) also requires a large number of transactions per quarter because the value of the characteristics (shadow price) is also estimated each period. It can be used with either micro or macro location variables, as the basket from period 0 is normally used as the reference basket. Like the characteristics method, it is very responsive to price changes. Nevertheless, the reference basket must also be updated regularly, if it is to remain representative of the real estate market, especially when it is dynamic. Its geometric formula with double imputation is[12]:

$$P_{HI} = \frac{\left(\prod_{i=1}^{N^0} \hat{\beta}_k^t (z_{ik}^0)\right)^{\frac{1}{N^0}}}{\left(\prod_{i=1}^{N^0} \hat{\beta}_k^0 (z_{ik}^0)\right)^{\frac{1}{N^0}}}$$

[10] Alternatively, it is possible to pool different quarters if there are few transactions. However, the index becomes less responsive, as the rolling time dummy hedonic method.
[11] For more information on the formulas of the characteristics prices approach please consult: Handbook on Residential Property Prices Indices (RPPIs), Eurostat 2013 or Econometric Issues in Hedonic Property Price Indices, Silver Mick 2022
[12] For more information on the formulas of the hedonic imputation approach please consult: Handbook on Residential Property Prices Indices (RPPIs), Eurostat 2013 or Econometric Issues in Hedonic Property Price Indices, Silver Mick 2022

$\hat{\beta}_k^t$ - *shadow price of characteristic k in period t*
$\hat{\beta}_k^0$ - *shadow price of characteristic k in period 0*
z_{ik}^0 - *characteristic k of object i in period 0*
N^0 - *number of objects in period 0*

When we have few transactions per quarter, like in Switzerland, with its average of 7'000 transactions per quarter, when we have geo-localized variables and when we want to stratify the sample in order to publish sub-indices, none of the above methods is appropriate, with the exception maybe of the rolling time dummy hedonic method. The following method may be the solution.

The hedonic repricing method is a variant of the imputation method. It has the advantage of using the actual prices and characteristics of both periods and correcting the impact of the different characteristics between the two periods on the result by using a shadow price of each characteristic. By applying the shadow prices to the current and previous periods, the qualitative (or quantitative) difference between the two samples is estimated. The shadow prices ($\hat{p}$) of the characteristics (z) are based on a pooling of 1 to 2 years of data (period b); this pooling allows to obtain stable and good quality coefficients, which must be updated periodically, especially if the real estate market is unstable. Otherwise, the quality of the index threatens to deteriorate continuously[13]. Its formula is:

$$P_{HR} = \frac{\left(\prod_{i=1}^{N^t} p_i^t\right)^{\frac{1}{N^t}}}{\left(\prod_{i=1}^{N^0} p_i^0\right)^{\frac{1}{N^0}}} \div \frac{\left(\prod_{i=1}^{N^t} \hat{\beta}_k^b (z_{ik}^t)\right)^{\frac{1}{N^t}}}{\left(\prod_{i=1}^{N^0} \hat{\beta}_k^b (z_{ik}^0)\right)^{\frac{1}{N^0}}}$$

p_i^t - *transaction price of object i in period t*
p_i^0 - *transaction price of object i in period 0*
$\hat{\beta}_k^b$ - *shadow price of characteristic k in period b*
z_{ik}^t - *characteristic k of object i in period t*
z_{ik}^0 - *characteristic k of object i in period 0*
N^t - *number of objects in period t*
N^0 - *number of objects in period 0*

Let's take a very simple example to illustrate the use of this method. The transaction data for the entire year T-1 was used to estimate the following linear hedonic model:

$$\mathrm{Y} = 50'000 + 100'000\mathrm{X}$$

Y - *predicted price of dwelling I*
X - *number of rooms*

[13] Hill Robert, Scholz Michael, Shimizu Chihiro, Steurer Miriam (2017); An evaluation of the hedonic methods used by European National Statistical Institutes to compute their official House Price Indices.

Let the following transaction data be collected during quarter 1 and 2 of year T

Quarter 1 Year T

	Rooms	*Price*
Dwelling A :	2	260'000.-
Dwelling B :	4	460'000.-
Geometric mean :		345'832.33

Quarter 2 Year T

	Rooms	*Price*
Dwelling C :	3	330'000.-
Dwelling D :	5	540'000.-
Dwelling E :	3	460'000.-
Geometric mean :		434'398.69

Quality unadjusted price index : **125.6096**

Applying the hedonic model to the transaction data yields the following predicted prices:

Quarter 1 Year T

	Rooms	*Predicted price*
Dwelling A :	2	250'000.-
Dwelling B :	4	450'000.-
Geometric mean :		335’410.20

Quarter 2 Year T

	Rooms	*Predicted price*
Dwelling C :	3	350'000.-
Dwelling D :	5	550'000.-
Dwelling E :	3	350'000.-
Geometric mean :		406’911.15

Quality adjustment index : **121.3175**

Quality adjusted price index : **103.5380**

By compiling the price change with transaction prices, we obtain an index of 125.6, thus an increase of 25.6% in quality unadjusted transaction prices. However, the two samples are different because the dwellings have different number of rooms. By calculating the quality adjustment index, we apply the same shadow price to the characteristic of the two samples, which gives a quality index (or a quantity index) of 121.3, which will be used to correct the quality unadjusted price index and give the quality adjusted price index of 103.5 thus a price increase corrected for qualitative differences of 3.5%.

The hedonic repricing is a method close to the imputation method because we can also transform its formula as follows:

$$P_{HR} = \frac{(\prod_{i=1}^{N^t} p_i^t)^{\frac{1}{N^t}}}{(\prod_{i=1}^{N^0} p_i^0)^{\frac{1}{N^0}}} \div \frac{(\prod_{i=1}^{N^t} \hat{\beta}_k^b(z_{ik}^t))^{\frac{1}{N^t}}}{(\prod_{i=1}^{N^0} \hat{\beta}_k^b(z_{ik}^0))^{\frac{1}{N^0}}}$$

$$P_{HR} = \frac{(\prod_{i=1}^{N^t} p_i^t)^{\frac{1}{N^t}}}{(\prod_{i=1}^{N^0} p_i^0)^{\frac{1}{N^0}}} \times \frac{(\prod_{i=1}^{N^0} \hat{\beta}_k^b(z_{ik}^0))^{\frac{1}{N^0}}}{(\prod_{i=1}^{N^t} \hat{\beta}_k^b(z_{ik}^t))^{\frac{1}{N^t}}}$$

$$P_{HR} = \frac{(\prod_{i=1}^{N^t} p_i^t)^{\frac{1}{N^t}}}{(\prod_{i=1}^{N^t} \hat{\beta}_k^b(z_{ik}^t))^{\frac{1}{N^t}}} \times \frac{(\prod_{i=1}^{N^0} \hat{\beta}_k^b(z_{ik}^0))^{\frac{1}{N^0}}}{(\prod_{i=1}^{N^0} p_i^0)^{\frac{1}{N^0}}}$$

$$P_{HR} = \left[\frac{\prod_{i=1}^{N^t} p_i^t}{\prod_{i=1}^{N^t} \hat{\beta}_k^b(z_{ik}^t)}\right]^{\frac{1}{N^t}} \times \left[\frac{\prod_{i=1}^{N^0} \hat{\beta}_k^b(z_{ik}^0)}{\prod_{i=1}^{N^0} p_i^0}\right]^{\frac{1}{N^0}}$$

and present the calculation table as follows:

	Price Q1	*Predicted price Q1*	*Price Q2*	*Predicted price Q2*	*Elementary index Q2 with sample Q1*	*Elementary index Q2 with sample Q2*
Dwelling A :	260'000.-			*250'000.-*	96.1538	
Dwelling B :	460'000.-			*450'000.-*	97.8261	
Dwelling C :		*350'000.-*	330'000.-			94.2857
Dwelling D :		*550'000.-*	540'000.-			98.1818
Dwelling E :		*350'000.-*	460'000.-			131.4286
Geometric mean:					96.9864	106.7552
Quality adjusted price index:						**103.5380**

For each raw transaction, a price is imputed to the missing period, using the shadow price of the characteristics of the hedonic model, and an elementary index is calculated for each transaction. The geometric mean of the elementary indices is calculated for the sample of quarter 1 and the sample of quarter 2. The two indices are then multiplied to obtain the quality adjusted price index.

3. The comparison between rolling time dummy and hedonic repricing

To ensure the adequacy of the hedonic repricing, we calculated the results using the rolling time dummy (rolling window of the previous 4 quarters), with two variants: without stratification and with stratification. The rolling time dummy index is not published, is used solely as an internal benchmark and to monitor the development of the implicit prices.

In the following figures, the hedonic repricing indices are compared with the stratified rolling time dummy indices.

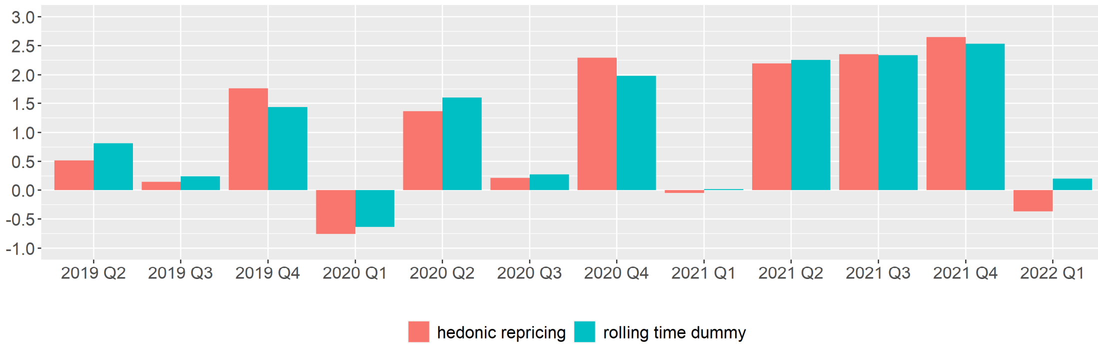

Figure 2: Total index – quarterly change rates (%) hedonic repricing vs. rolling time dummy

For the total property price indices, the results of the hedonic repricing and the rolling time dummy method appear to be relatively strongly correlated. However, there are a few quarters where the divergences are somewhat larger. Particularly noteworthy is the first quarter 2022 where the two indices point in opposite directions. The rolling time dummy index appears to be a bit smoother than the hedonic repricing index.

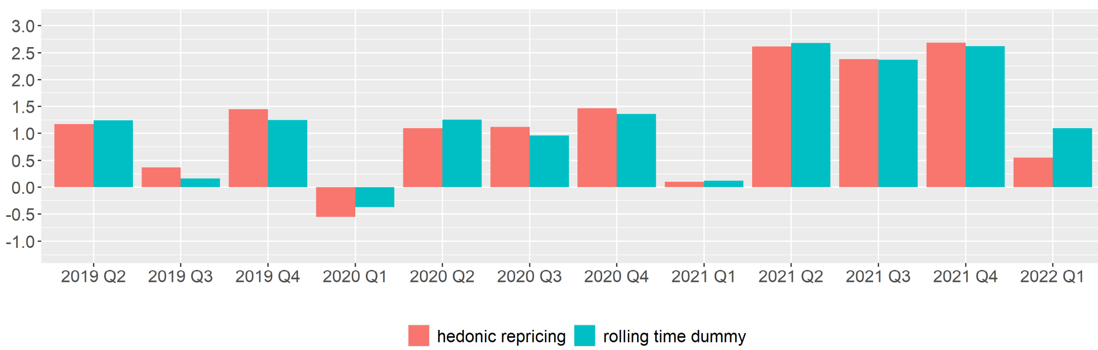

Figure 3: Single-family-house index – quarterly change rates (%) hedonic repricing vs. rolling time dummy

For the single-family houses, the hedonic repricing index also shows a slightly larger variance than the rolling time dummy index. The most significant difference between the two indices is again to be found in the first quarter 2022. In the other quarters, the differences are rather small.

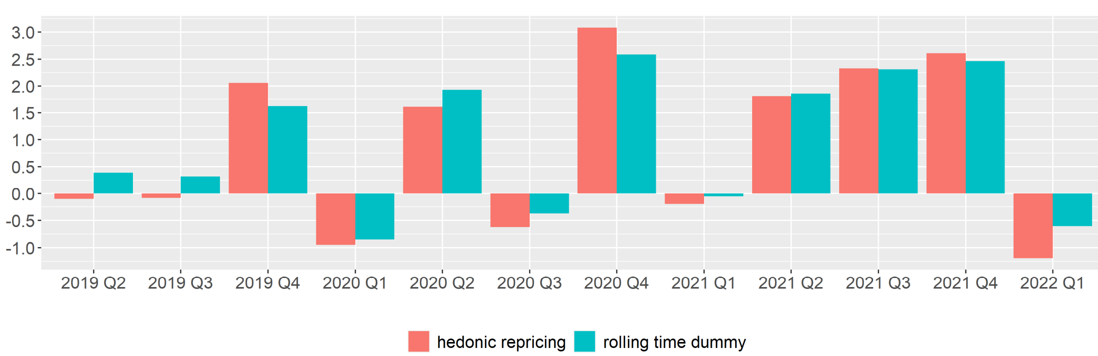

Figure 4: Condominium index – quarterly change rates (%) hedonic repricing vs. rolling time dummy

The pattern for the overall and the single-family house indices is also mirrored in the index for condominiums. Although, the quarterly differences between the hedonic repricing index and the rolling time dummy index are slightly more pronounced for condominiums than for single-family houses. In the second and third quarter 2019, the two indices point in opposite directions.

4. The model update and its impact on the results

The Swiss hedonic model is based on about 60'000 transaction data covering 2 years (T-2 and T-1). It is updated every year. Since the Swiss residential property price index is still very young, there have only been two updates so far: one for the published results of 2021 and one for the published results of 2022 (only one quarter available).

If the model had not been updated, the results would have been very slightly different, as the graphs below show. The first model update took place in the first quarter 2021. Therefore, the differences between the indices are only visible from this point onwards. That is why only the specific time periods are shown in the figures below.

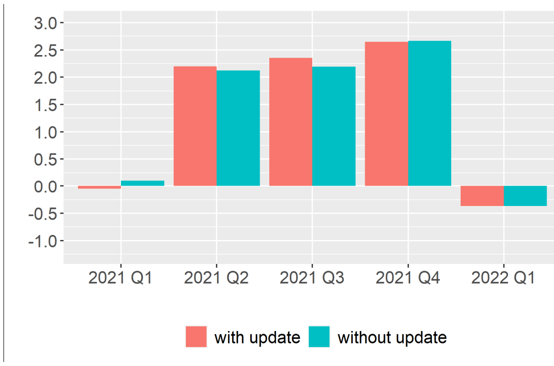

Figure 5: Total index – quarterly change rates (%) hedonic repricing with and without updates

For the residential property, the indices with and without model updates show only small differences. What is most striking is that the indices somewhat point in different directions in the first quarter of 2021.

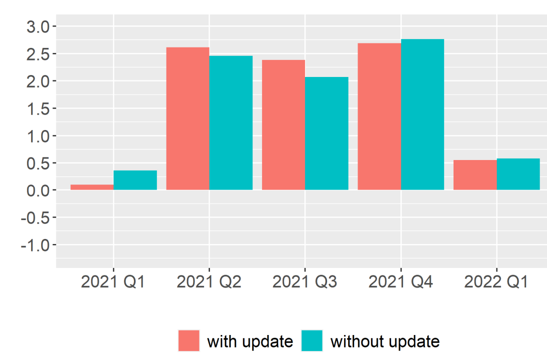

Figure 6: Single-family-house index – quarterly change rates (%) hedonic repricing with and without updates

Also for the single-family houses, minor differences between the indices with and without model updates can be seen.

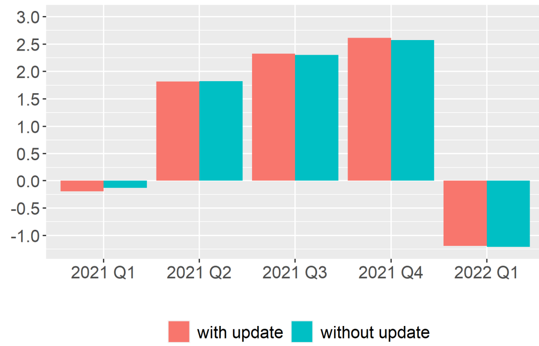

Figure 7: Condominium index – quarterly change rates (%) hedonic repricing with and without updates

For the condominiums, the differences between the indices with and without model updates are the narrowest.

Even if the model update has a small impact on the results, it is still visible. The annual model update is therefore certainly justified. The model update is mainly necessary to integrate new characteristics that have a determining impact on the price.

5. The decomposition of the quality effect and the price effect

The hedonic repricing method has the advantage of being able to directly decompose the price effect and the quality effect. The figures below show the gross, the quality and the quality-adjusted index.

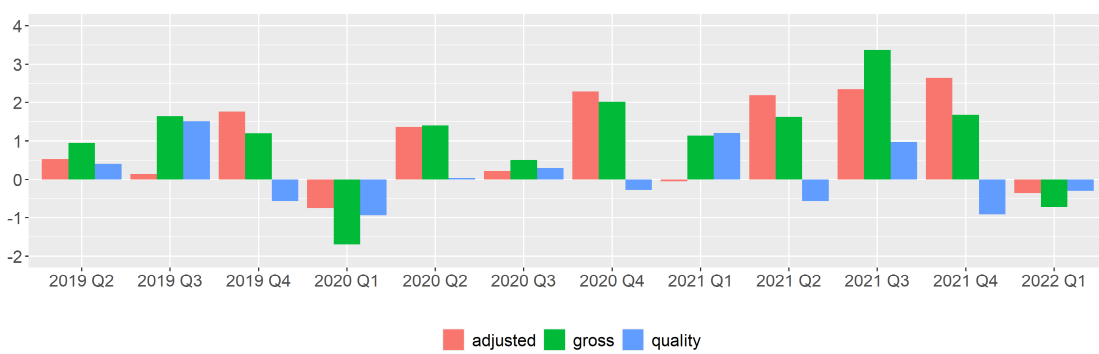

Figure 8: Total index – quarterly change rates (%) gross, quality and adjusted index

In the case of the overall index for residential property, it becomes clear that the results would look significantly different without quality adjustment.

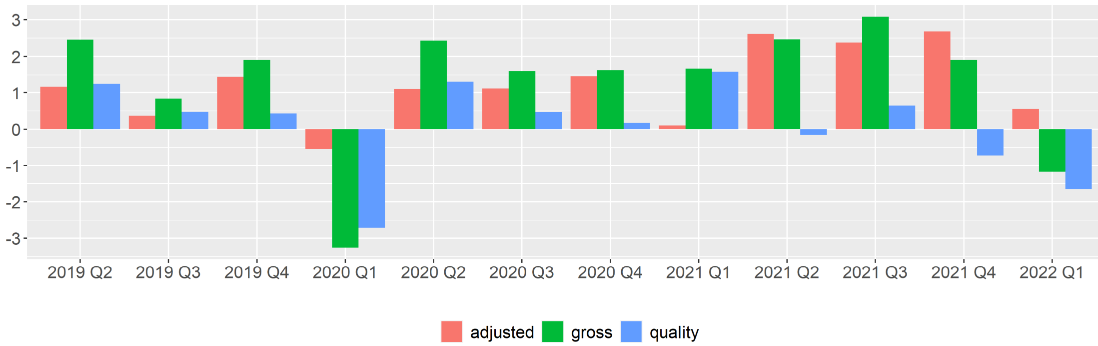

Figure 9: Single-family-house index - quarterly change rates (%) gross, quality and adjusted index

A similar picture emerges for the single-family houses. In some quarters, such as the first quarter 2020, the influence of the quality adjustment procedure is particularly evident.

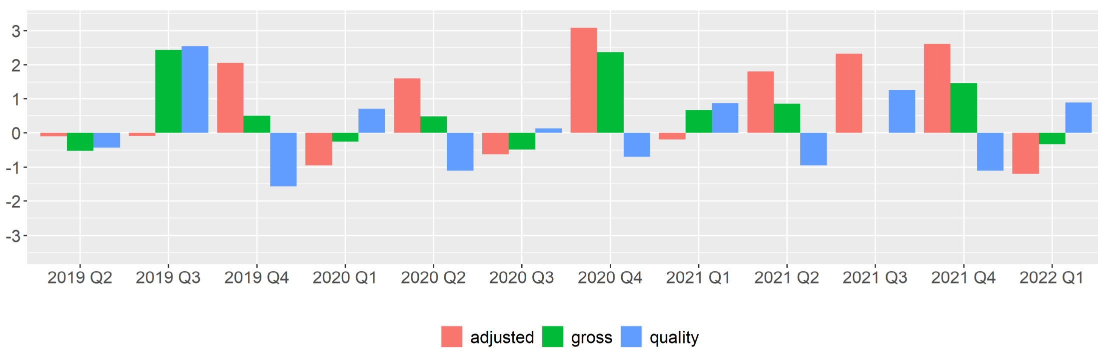

Figure 10: Condominium index - quarterly change rates (%) gross, quality and adjusted index

In the case of the condominiums, too, there are quarters in which the influence of the quality adjustment is pronounced. This is for example the case in the third quarter 2019.

The quality adjustment has a very strong impact on the quarterly and annual evolutions, as shown in the quarterly evolutions above. Without this qualitative adjustment, the quarterly price evolutions would have been very different.

6. Rolling hedonic repricing

The main criticism of the hedonic repricing method is the lazy impulse it suggests since the model can remain fixed. What if we updated the model every quarter using the last 8 quarters of data? We could call this approach rolling hedonic repricing.

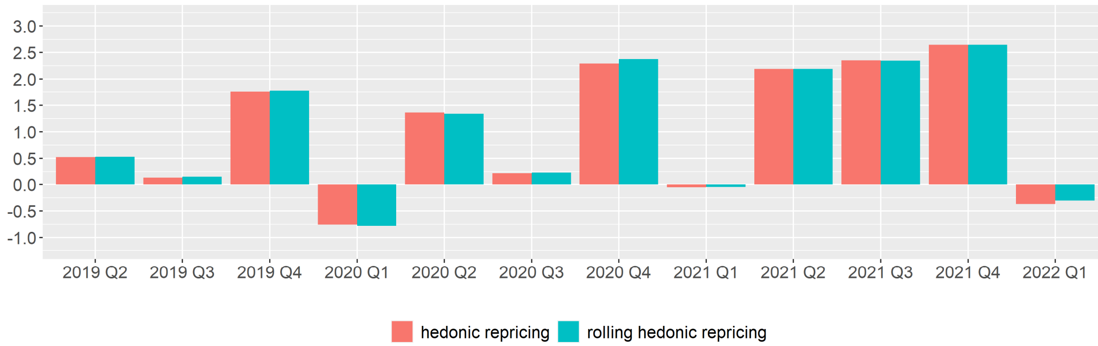

Figure 11: Total index – quarterly change rates (%) hedonic repricing vs. rolling hedonic repricing

For the total residential property, quarterly change rates of the hedonic repricing index and the rolling hedonic repricing are basically identical.

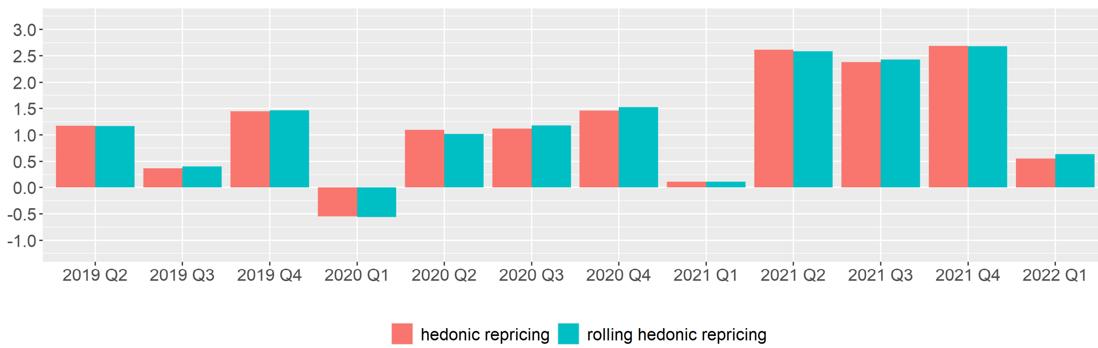

Figure 12: Single-family-house index – quarterly change rates (%) hedonic repricing vs. rolling hedonic repricing

The same applies to the object type of single-family houses, where the indices for both approaches behave very similarly.

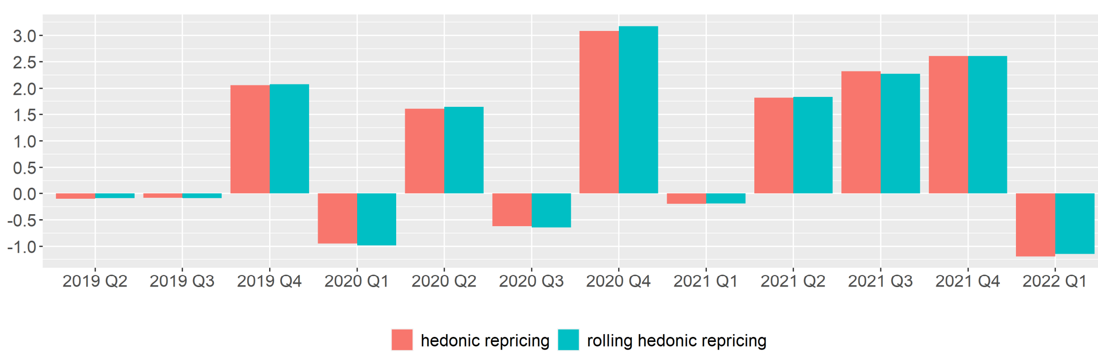

Figure 13: Condominium index – quarterly change rates (%) hedonic repricing vs. rolling hedonic repricing

Last but not least, the hedonic repricing index and the rolling hedonic repricing index are also very similar for condominiums.

The results show hardly any differences between the two indices. Even though the available time series are still very short, we assume that an annual update of the hedonic models is sufficient. The quarterly updating of the hedonic models would be associated with a certain amount of work. In addition, chaining would have to be carried out every quarter.

7. Conclusions and way forward

The hedonic repricing method is appropriate when the number of transactions per period is low. It allows decomposing directly the price effect from the quality effect. It uses the real prices and characteristics of the two comparison periods. It corrects for qualitative differences in the two samples by applying the same shadow prices to both periods. These shadow prices must be updated at regular intervals, depending on the market dynamics. In our case, an annual update seems sufficient.

In addition, the indices calculated with the hedonic repricing method seems to be more responsive than the rolling time dummy price index. The data pooling from several periods seems to result in the rolling time dummy index being smoothed to a certain extent.

If we classify the hedonic methods according to the number of transactions and the number of characteristics available, as Residential property price index (RPPI) practical compilation guide, IMF 2020 has done, we could add the hedonic repricing method as a method in its own right and not as a method derived from the imputation method.

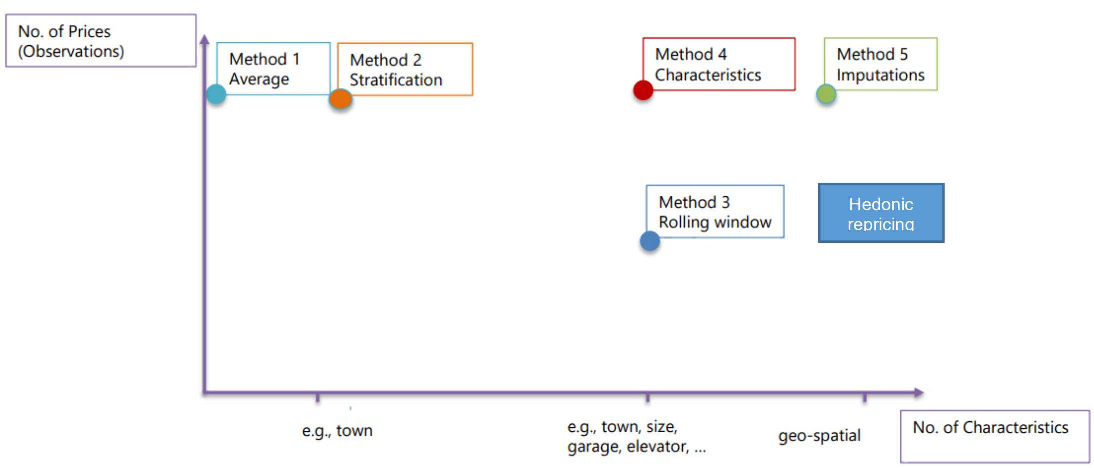

Figure 14: Classification scheme for quality adjustment methods

Finally, it should be noted that the hedonic repricing method is very simple to apply once the hedonic model is estimated. There is only one update of the model per year and only one chaining. It is also quite easy to explain to the public.

Apart from the choice of the hedonic method, it is important from the FSO's point of view to compare the approach used with others. There may be other possibilities for simulation: applying the imputation and characteristics method and pooling the data from 4 to 8 quarters in order to have good quality shadow prices. This will be the next step.

8. Table of references

Eurostat (2013); Handbook on Residential Property Prices Indices (RPPIs).

Hill Robert, Scholz Michael, Shimizu Chihiro, Steurer Miriam (2017); An evaluation of the hedonic methods used by European National Statistical Institutes to compute their official House Price Indices.

IMF (2020); Residential property price index (RPPI) practical compilation guide.

Silver Mick (2022); Econometric Issues in Hedonic Property Price Indices: Some Practical Help.

Silver Mick (2018); How to measure hedonic property price indexes better.

Swiss Federal Statistical Office (2020); Swiss Residential Property Price Index. Methodological Principles.

Swiss Federal Statistical Office (2020); Occupancy Status Statistics.