Fees with multiple price-dependent rates

Claude Lamboray *

1 Introduction

The development of an owner-occupied housing price index system includes measuring the additional costs related to the acquisition of a dwelling. The classification used in the European Owner-occupied housing regulation explicitly contains a heading covering "other services related to the purchase of a dwelling". Examples of such transaction costs are registration fees, fees for legal services or commissions paid to real estate agencies. These fees are typically defined as a percentage of the transaction price.

The treatment of a price that is proportional to a unit value is not new. The EU regulation No 1920/2001 on the treatment of service charges proportional to transaction values states that the price taken into account is the proportion multiplied by the value of a representative unit transaction in the base period. Changes in the rules defining the service charges should be reflected in the price index. Moreover, changes in the value of the representative unit transactions should also be shown as price change. This change can be estimated by using an appropriate price index. A similar issue can be found in regulation No 2166/1999 on income-dependent prices. This regulation states for instance that changes in the purchaser prices resulting from changes in purchasers' incomes shall be shown as price changes in the HICP.

It happens that the percentage applied to the transaction price is depending on the transaction price. For instance, taxes or other fees related to the acquisition of a dwelling can be progressive. This means that the higher the transaction price, the higher the rate. In practice, thresholds (upper and lower bounds) are defined. A transaction price can be assigned to a class if the price lies between an upper and a lower bound. Each class has a particular rate that is then applied to the transaction price.

The compilation of a price index for such a scheme with multiple classes can reveal an unstable behavior. Consider the following situation (see Table 1). We suppose that the rate is 1% if the transaction price of a dwelling is less than 350000 € and the rate is 4% if the transaction price of a dwelling is above 350000 €. In other words, there are two classes that are separated by a threshold value of 350000 €. For example, if the price of a dwelling is equal to 300000 €, then the corresponding fee is $1\% \cdot 300000 € = 3000 €$. We now assume that the price of this same dwelling changes to 310000 €, which is a 3% increase. The fee then increases at the same rate.

We now consider a dwelling (see Table 2) that is sold for a price of 345000 €. Consequently, the fee amounts to 3450 € as the transaction price is still below the threshold. Let us suppose that this same dwelling has a price of 355000 € in the following period. In such a situation house

*claude.lamboray@statec.etat.lu, STATEC, 13, rue Erasme, B.P.304, L-2013 Luxembourg

	House price	Rate applied	Fee
Base period	300 000	1%	3 000
Current period	310 000	1%	3 100
Variation	3%	0%	3%

Table 1: An example where the fee increases at the same rate than house prices.

	House price	Rate applied	Fee
Base period	345 000	1%	3 450
Current period	355 000	4%	14 200
Variation	3%	300%	312%

Table 2: An example where the fee increases at a much higher rate than house prices.

prices have gone up by 3% while at the same time transaction costs have risen by 312%.

In this small example, both the rates (1% and 4%) and the threshold value (350000 €) have remained unchanged. Although in both cases, house prices have gone up by 3%, the amount of the fee can increase either by 3% or by 312%, depending on the level of the price of the dwelling in the base period.

In the next section, we are going to formalize a price index for fees that are defined as multiple price dependent rates. In section 3, index formulas are provided assuming a log-normal distribution of house prices. The case of multiple price-dependent fixed fees is briefly addressed in section 4. An application of these index formulas is given in section 5. The main conclusions are summarized in section 6.

2 Conceptual framework

Our aim here is to measure fees defined as multiple price-dependent rates with a Laspeyres-type price index. Consequently, the characteristics of the service included in the basket needs to be somehow kept constant throughout time. To formalize this idea, we use the following notations. Let $F_t(P_t(X))$ be the total fee collected in t for a set of house prices, denoted by $P_t(X)$, where t is the reference period of the prices and X represents the set of houses to which these prices refer.

$$P_t(X) = \{p_t^i : i \in X\}.$$

A first idea, referred to as the "base line-solution" (see [3]), is to compare the fee collected in two periods for a given set of representative dwelling prices:

$$I_t = \frac{F_t(P_0(X))}{F_0(P_0(X))}.$$

This price index takes into account changes in the rule of the fee scheme, comparing the fee collected in the current period to the fee collected in the base period for a fixed set of transaction prices. However, this solution is not totally satisfactory with respect to EU regulation as it does not take into account changes in the unit transaction price. The following price index, called the "deflator solution" (see [3],[4]), takes into account both changes in the rule and changes in house prices.

$$I_t = \frac{F_t(P_t(X))}{F_0(P_0(X))}.$$

In this price index, the fee collected in the current period using current prices is compared to the fee collected in the base period using base prices. In order to be in line with the fixed-basket idea, the prices in the current period and in the base period refer to a set of dwellings X with the same characteristics. In other words, this is a kind of unit value price index with a reference stock of dwellings that is being priced at the two comparison periods.

In the simple case, the fee is defined as a single percentage r_t of the transaction price, whatever the transaction price. The deflator solution then reduces to the following:

$$I_t = \frac{r_t \cdot \sum_{p \in P_t(X)} p}{r_0 \cdot \sum_{p \in P_0(X)} p} = \frac{r_t}{r_0} \cdot HPI_t,$$

where HPI_t is a quality-adjusted house price index. In practice, the compilation of such a price index I_t is straight-forward. The percentage needs to be followed over time. This ratio is then multiplied by an appropriate housing price index.

In the case of multiple price-dependent rates, we denote by $P_t^s(X)$ the set of prices that belong at period t to the price class s defined by a lower bound l_t^s and an upper bound u_t^s. Every price that lies between these two boundaries is assigned to this price class.

$$P_t^s(X) = \{p_t^i : i \in X \text{ and } l_t^s \leq p_t^i < u_t^s\}.$$

Note that for the "first" price class, the lower bound is set to zero and for the "last" price class, there is no upper bound. Moreover, the upper bound of a class is the lower bound of the subsequent class.

Let r_t^s be the rate applied in price class s at period t. The deflator solution can then be disaggregated using the various price classes:

$$I_t = \frac{\sum_s r_t^s \cdot \sum_{p \in P_t^s(X)} p}{\sum_s r_0^s \cdot \sum_{p \in P_0^s(X)} p} = \sum_s \frac{r_0^s \cdot \sum_{p \in P_0^s(X)} p}{\sum_{s'} r_0^{s'} \cdot \sum_{p \in P_0^{s'}(X)} p} \cdot \frac{r_t^s}{r_0^s} \cdot \frac{\sum_{p \in P_t^s(X)} p}{\sum_{p \in P_0^s(X)} p}.$$

The price index I_t can be decomposed into the contributions of each price class. For each price class, there are three ratios that are being multiplied. In the first term, the fee collected in the base period in price class s is compared to the total fee collected in the base period. Hence this represents the weight of this class.

$$w^s = \frac{r_0^s \cdot \sum_{p \in P_0^s(X)} p}{\sum_{s'} r_0^{s'} \cdot \sum_{p \in P_0^{s'}(X)} p}.$$

The second term is a simple ratio of class-specific rates. The last term is a comparison of house prices restricted to price class s. We call this a "class-confined house price index" :

$$HPI_t^s = \frac{\sum_{p \in P_t^s(X)} p}{\sum_{p \in P_0^s(X)} p}.$$

Note that this class-confined house price index is in general different from the overall house price index, except in the degenerated case where there is only a single price class. In fact, the number of prices in the numerator and in the denominator of this index is not necessarily the same. This will be illustrated in the example at the end of this section. However, the number of prices over all price classes is kept constant:

$$\sum_s |P_t^s(X)| = \sum_s |P_0^s(X)| = |X|.$$

In the end, the deflator solution can thus be rewritten as follows:

$$I_t = \sum_s w^s \cdot \frac{r_t^s}{r_0^s} \cdot HPI_t^s.$$

To compile such a price index, a solution consists in adjusting a set of representative house prices of the base period using an overall house price change. In other words, the transactions in the base period are used to define a reference stock of dwellings X. The price of these reference dwellings is then price updated using the HPI_t.

$$p_t^i = p_0^i \cdot HPI_t \quad \forall i \in X.$$

This approach is of course an approximation because it assumes that the price change is independent of the price level. It makes it nevertheless possible to assign prices prices to a particular price class both in the current period and in the base period. Consequently, the price index for the fee can be estimated.

To illustrate this, let us consider an example with three price classes. In the first class (0 € - 350000 €), a rate of 1% is applied, in the second class (350000 € - 500000 €) a rate of 4% is applied and in the third class ($\geq$ 500 000 €) a rate of 7% is applied. We suppose that these rates are increased by 0.5 percentage points to 1.5% (1^{st} price class), 4.5% (2^{nd} price class) and 7.5% (3^{rd} price class). There are 7 house prices in the base period and we suppose that the overall price change is 10% between the base period and the current period. Table 3 summarizes this situation.

In this example, the three class-confined house price indices are then:

$$\begin{aligned}
HPI_t^1 &= \frac{220000 + 330000}{200000 + 300000 + 340000} = 0.65 \\
HPI_t^2 &= \frac{374000 + 495000}{450000 + 480000} = 0.93 \\
HPI_t^3 &= \frac{528000 + 660000 + 770000}{600000 + 700000} = 1.51
\end{aligned}$$

Note that for instance the first price class contains 3 prices in the base period but only 2 prices in the current period. In fact, the 10% price increase makes house price number 3 move

<table>
<tr><th></th><th>Prices in the base period</th><th>Rate in the base period</th><th>Prices in the current period: +10%</th><th>Rate in the current period</th></tr>
<tr><td>House price 1</td><td>200000 €</td><td rowspan="3">1%</td><td>220000 €</td><td rowspan="2">1.5%</td></tr>
<tr><td>House price 2</td><td>300000 €</td><td>330000 €</td></tr>
<tr><td>House price 3</td><td>340000 €</td><td>374000 €</td><td rowspan="2">4.5%</td></tr>
<tr><td>House price 4</td><td>450000 €</td><td rowspan="2">4%</td><td>495000 €</td></tr>
<tr><td>House price 5</td><td>480000 €</td><td>528000 €</td><td rowspan="3">7.5%</td></tr>
<tr><td>House price 6</td><td>600000 €</td><td rowspan="2">7%</td><td>660000 €</td></tr>
<tr><td>House price 7</td><td>700000 €</td><td>770000 €</td></tr>
</table>

Table 3: An example with 7 house prices and 3 price classes.

to the second price class.

Finally, the weights correspond to the share of the fee collected in each price class in the base period:

$$\begin{aligned}
TOT &= 1\% \cdot (200000 + 300000 + 340000) + 4\% \cdot (450000 + 480000) \\
&\quad +7\% \cdot (600000 + 700000) \\
&= 136600 \\
w^1 &= \frac{1\% \cdot (200000 + 300000 + 340000)}{136600} \\
&= 6.1\% \\
w^2 &= \frac{4\% \cdot (450000 + 480000)}{136600} \\
&= 27.2\% \\
w^3 &= \frac{7\% \cdot (600000 + 700000)}{136600} \\
&= 66.6\%
\end{aligned}$$

Consequently, the final index is obtained as follows:

$$I_t = 6.1\% \cdot \frac{1.5\%}{1.0\%} \cdot 0.65 + 27.2\% \cdot \frac{4.5\%}{4.0\%} \cdot 0.93 + 66.6\% \cdot \frac{7.5\%}{7.0\%} \cdot 1.51 = 1.42$$

In practice, if individual price data is available, then these prices could be index-linked and the price index for the fee scheme can be naturally compiled as shown in the preceding example. In some situations, there is no individual house price data available underlying the fee scheme that we wish to analyze. However, there is maybe an alternative source that provides information on the distribution of house prices. For instance, tax authorities sometimes disseminate transaction data by price class. Such information can be exploited to define a "representative" set of house prices for the base period.

3 Compilation of a price index assuming a log-normal distribution of house prices

A typical assumption made in empirical analysis on house prices is that they follow a log-normal distribution. This means that the logarithm of house prices is normally distributed. Recently, this assumption has been criticized (see [6], [5]), with some empirical evidence that house prices may have fatter tails than a log-normal distribution. In our context, the aim is not to measure with precision the distribution of house prices, but to compile a price index for fees based on a "representative" stock of house prices. For this purpose the log-normal approximation may be sufficient.

A log-normal distribution can be characterized with two parameters m (location parameter) and s (scale parameter). The first parameter relates to the mean and the second parameter to the standard deviation of the logarithm of the price. From an operational point of view, the parameters of the log-normal distribution can be reasonably well obtained. The reference period used to estimate the distribution should of course be related to the base period of the price index for the fee. However, this reference period could also be handled in a more flexible way if, for instance, the number of transactions is insufficient to estimate robust parameters.

In the preceding section, we have defined class-confined house price indices as a ratio of a sum of those house prices that fit into a given price class. Instead of actual price data, we now assume that P_t^i and P_0^i are two random variables. We are looking at the expected value of the ratio based on these random variables rather than generating random prices. It can then be shown that (see Appendix 1):

$$\begin{aligned} HPI_t^s &\equiv E\left[\frac{\sum_{l_t^s \leq P_t^i < u_t^s} P_t^i}{\sum_{l_0^s \leq P_0^i < u_0^s} P_0^i}\right] \\ &\approx \frac{E[P_t^i | l_t^s \leq P_t^i < u_t^s] \cdot P(P_t^i | l_t^s \leq P_t^i < u_t^s)}{E[P_0^i | l_0^s \leq P_0^i < u_0^s] \cdot P(P_0^i | l_0^s \leq P_0^i < u_0^s)} \end{aligned}$$

A practical result of a log-normal distribution is that partial expectations can be evaluated using the cumulative distribution function of the standard normal distribution, denoted by ϕ . If X is log-normally distributed with parameters m and s, it is known that (see Appendix 2 for a proof):

$$E[X|X \geq k] \cdot P(X|X \geq k) = e^{m+\frac{1}{2}s^2} \cdot \phi\left(\frac{m + s^2 - ln(k)}{s}\right).$$

In our setting, we suppose that transaction prices in the base period are log-normally distributed with parameters m_0 and s_0. In period t, prices have increased by a ratio given by the house price index HPI_t. It can then be shown that transaction prices in period t are log-normally distributed with parameters $m_0 + ln(HPI_t)$ and s_0. Exploiting the result on the conditional expectations of a log-normal distribution, the following index is obtained after some algebra :

$$HPI_t^s \approx HPI_t \cdot \frac{E_t^s}{E_0^s},$$

with:

$$E_t^s = \phi\left(\frac{m_0 + ln(HPI_t) + s_0^2 - ln(l_t^s)}{s_0}\right) - \phi\left(\frac{m_0 + ln(HPI_t) + s_0^2 - ln(u_t^s)}{s_0}\right)$$

For the "first" price class, the lower bound is set to 0 and the formula is as follows:

$$E_t^s = 1 - \phi\left(\frac{m_0 + ln(HPI_t) + s_0^2 - ln(u_t^s)}{s_0}\right).$$

For the "last" price class, there is no upper bound and the formula reduces to:

$$E_t^s = \phi\left(\frac{m_0 + ln(HPI_t) + s_0^2 - ln(l_t^s)}{s_0}\right).$$

The weights representing the relative expenditure of the fee in the base period can also be defined in terms of the underlying lognormal distribution:

$$\begin{aligned} w^s &\equiv E[\frac{r_0^s \sum_{l_0^s \leq P_0^i < u_0^s} P_0^i}{\sum_{s'} r_0^{s'} \sum_{l_0^{s'} \leq P_0^i < u_0^{s'}} P_0^i}] \\ &\approx \frac{r_0^s \cdot E_0^s}{\sum_{s'} r_0^{s'} \cdot E_0^{s'}}. \end{aligned}$$

The price index for a fee with multiple price-dependant rates can thus be compiled in practice as follows, assuming a log-normal distribution of prices in the base period:

$$I_t = \sum_s w^s \cdot \frac{r_t^s}{r_0^s} \cdot HPI_t^s \approx \sum_s \frac{r_0^s \cdot E_0^s}{\sum_s' r_0^{s'} \cdot E_0^{s'}} \cdot \frac{r_t^s}{r_0^s} \cdot \frac{E_t^s}{E_0^s} \cdot HPI_t.$$

Note that under a constant fee scheme, the price index for the fee I_t is not equal to the house price index HPI_t because in general $E_t^s \neq E_0^s$. An equivalence between the price index for the fee and the house price index can be obtained if for instance the thresholds of the price classes are adjusted with the overall house price change, i.e. $u_t^s = u_0^s \cdot HPI_t$ and $l_t^s = l_0^s \cdot HPI_t$.

4 The case of multiple price-dependent fixed fees

It may also happen that the price is defined as a fixed fee and not as a percentage of a transaction price. Moreover there can be multiple fixed fees depending on the level of the transaction price. The framework described in the previous two sections can also be applied to such a situation.

More formally, let us suppose that at period t the fee amounts to f_t^s if the transaction price belongs to price class s. In such a setting, a price index for the fee can be disaggregated into the different price classes as follows:

$$I_t = \sum_s \tilde{w}^s \cdot \frac{f_t^s}{f_0^s} \cdot \widetilde{HPI}_t^s,$$

where:

$$\tilde{w}^s = \frac{f_0^s |P_0^s(X)|}{\sum_{s'} f_0^{s'} |P_0^s(X)|} \qquad \widetilde{HPI}_t^s = \frac{|P_t^s(X)|}{|P_0^s(X)|}.$$

We may assume a log-normal distribution on the house prices in the base period with parameters m_0 and s_0. With similar considerations than in section 3, class-confined house price indices can be estimated in terms of the underlying distribution of house prices. Again, we use the cumulative distribution function of the standard normal distribution ϕ:

$$\widetilde{HPI}_t^s \approx \frac{\tilde{E}_t^s}{\tilde{E}_0^s},$$

where:

$$\tilde{E}_t^s = \phi\left(\frac{ln(u_t^s) - m_0 - ln(HPI_t)}{s_0}\right) - \phi\left(\frac{ln(l_t^s) - m_0 - ln(HPI_t)}{s_0}\right).$$

These formulas have to be adjusted to cope with the first class (lower bound is set to 0) and the last class (there is no upper bound).

The weights naturally follow the same line of reasoning:

$$\tilde{w}^s \approx \frac{f_0^s \cdot \tilde{E}_0^s}{\sum_{s'} f_0^{s'} \cdot \tilde{E}_0^{s'}}.$$

5 Example: The registration fee in Luxembourg

We now illustrate the index formulas on real data from Luxembourg. As an overall price change, we use the house price index published by STATEC. The transaction prices of dwellings (apartments and single-family houses) recorded in 2007 can be approximated with a log-normal distribution (see Figure 1). Using the maximum likelihood estimation, the parameters are $m_0 = 12.69$ and $s_0 = 0.44$.

In principle, the registration fee for the acquisition of a dwelling amounts to 7% of the transaction price. In 2002, a tax allowance of 20000 € on these registration fees has been introduced for households acquiring a dwelling. This means that registration fees up to 20000 € are neutralized. In all case, the buyer must pay at least 100 €. In other words, if 7% times the transaction price is less than 20100 € (i.e. that the transaction price is less than $\frac{20100}{0.07}$ =287143 €), then the fee amounts to 100 €, otherwise the fee amounts to 7% times the transaction price less 20000 €. Consequently, this fee scheme has two classes: in the first price class a fixed fee has to be paid, in the second price class a rate is applied to the transaction price:

$$r(p) = \begin{cases} 100 & \text{if } p \leq 287143 \\ 7\% \cdot p - 20000 & \text{if } p > 287143 \end{cases}$$

It is possible to combine the index formulas introduced in sections 3 and 4 to cope with both a fixed fee (the first price class) and a rate (the second price class). Moreover the tax allowance reduces the effective rate applied. We will thus compute an implicit rate depending on how prices are distributed in the second price class in the respective periods. The price index for such a fee can thus be disaggregated into the two price classes (the detailed index formulas can be found in Appendix 3):

$$I_t = \widetilde{w}^1 \cdot \frac{f_t^1}{f_0^1} \cdot \widetilde{HPI}_t^1 + w^2 \cdot \frac{r_t^2}{r_0^2} \cdot HPI_t^2$$

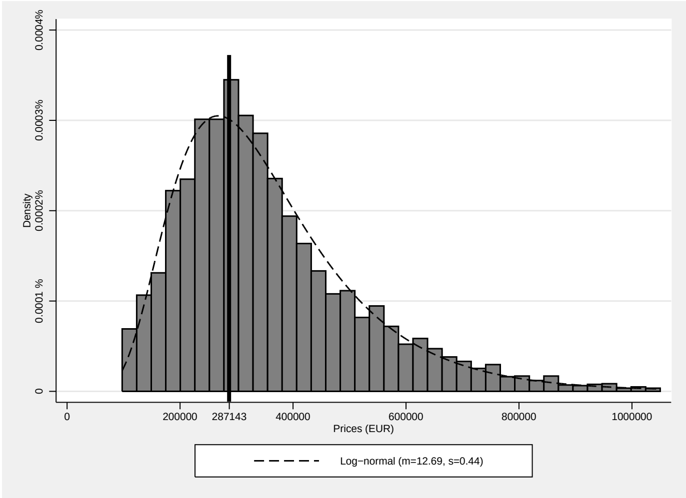

Figure 1: Distribution of transaction prices (apartments and single-family houses) in Luxembourg in 2007.

First, we suppose that the fee scheme is kept constant over time. This means that the price index for this fee is driven only by the overall price change in transaction prices measured by the house price index. Looking at the results (see Figure 2), we can see that the price index for the fee correlates with the house price index: if one index moves up (or moves down), then the other index also moves up (or moves down). However, the price index for the fee is much more volatile. On average the quarterly rate of change for the fee price index is 2.7 times the quarterly rate of change for the house price index. At the end of the simulation (2^{nd} quarter 2011), the index measuring the fee stands at 1.26 points, whereas the house price index only reaches 1.09 points.

The detailed results of all the parameters of the index formula can be found in Table 4. The weight for the first price class is insignificant compared to that of the second class. In fact, there is almost no fee to pay if a transaction is in the first price class. However, this does not mean that we can simply ignore the first price class. It also matters how prices are distributed in the second price class and how they move from the first price class to the second price class. The implicit rate of the second price class is not 7%, but rather something between 2.5% and 2.7%. The location parameter of the log-normal distribution increases over time. Consequently more and more prices are above the 287143 € threshold, which results in a sharp decrease of the first class-confined price index whereas at the same time the second class-confined price index is increasing.

A different price distribution in the base period leads to a different price index for the registration fees. We assumed in this example that the parameters of the log-normal distribution in the base period are $m_0 = 12.69$ and $s_0 = 0.44$. We test for the sensitivity of these parameters on the results. The location parameter is increased or decreased by 0.1, which roughly corresponds to an increase or a decrease of 10% of the average transaction price. The scale parameter is increased or decreased by 0.5. Combining all the possibilities we have simulated 8 alternative

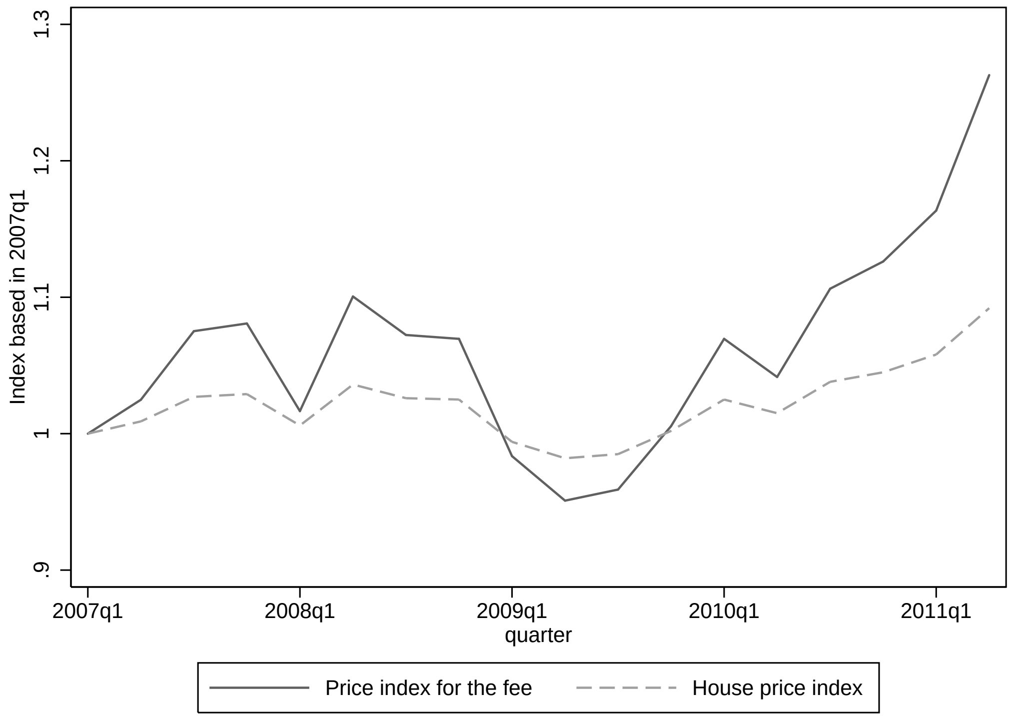

Figure 2: The price index for the registration fee adjusted with the house price index.

	Fee	$\tilde{w}^1$	f^1	HPI^1	w^2	r^2	HPI^2
2007 01	1.000	0.005	100	1.000	0.995	2.53%	1.000
2007 02	1.025	0.005	100	0.980	0.995	2.55%	1.017
2007 03	1.075	0.005	100	0.941	0.995	2.59%	1.052
2007 04	1.081	0.005	100	0.937	0.995	2.60%	1.055
2008 01	1.017	0.005	100	0.987	0.995	2.55%	1.011
2008 02	1.101	0.005	100	0.922	0.995	2.61%	1.069
2008 03	1.072	0.005	100	0.943	0.995	2.59%	1.050
2008 04	1.070	0.005	100	0.945	0.995	2.59%	1.048
2009 01	0.984	0.005	100	1.013	0.995	2.52%	0.988
2009 02	0.951	0.005	100	1.041	0.995	2.50%	0.965
2009 03	0.959	0.005	100	1.034	0.995	2.50%	0.971
2009 04	1.006	0.005	100	0.996	0.995	2.54%	1.004
2010 01	1.070	0.005	100	0.945	0.995	2.59%	1.048
2010 02	1.042	0.005	100	0.967	0.995	2.57%	1.029
2010 03	1.106	0.005	100	0.918	0.995	2.62%	1.072
2010 04	1.126	0.005	100	0.903	0.995	2.63%	1.086
2011 01	1.164	0.005	100	0.877	0.995	2.66%	1.110
2011 02	1.263	0.005	100	0.810	0.995	2.73%	1.174

Table 4: The parameters of the price index for the fee.

	m_0	$m_0 + 0.1$	$m_0 - 0.1$
s_0	0.000	-0.005	0.006
$s_0 + 0.05$	-0.005	-0.009	0.000
$s_0 - 0.05$	0.006	-0.001	0.014

Table 5: : Average difference in index points between the base scenario and alternative scenarios.

scenarios that can be compared to the base-line solution (see Table 5). Overall, the results are quite robust in the sense that the average difference between the base-line solution and the alternative scenarios never exceeds 0.02 index point. Changes in the location parameter have less impact than changes in the scale parameter.

There had been plans by the Luxembourg Government at the end of 2010 to reduce the tax allowance in the scheme from 20000 € to 10000 €. Such a change in the rule of the scheme can be easily captured by the price index. In fact, the threshold value is reduced and the compilation of the implicit rates has to be adjusted accordingly. If this policy change had been implemented at the beginning of 2011, the rate of change between 4Q2010 and 1Q2011 of the price index would have jumped to 110.8%, against 3.3% under constant policy conditions. House prices driving the underlying distribution have increased by 1.2% between these two quarters.

In this example, we have kept the base period (1^{st} quarter 2007) constant over several years in order to illustrate the index formulas. An alternative approach consists in re estimating each year the parameters of the distribution and compiling an annually chained price index.

6 Conclusion

The compilation of a price index reflecting a fee proportional to a transaction value is typically treated by comparing the proportion factor over time, multiplied by an appropriate price index. In this paper we have generalized this type of expenditure to a fee with multiple price-dependent rates. Such a scheme can be encountered in the context of transaction costs associated with the acquisition of a dwelling.

The price index for such a fee can be naturally disaggregated into various sub-components. In each price class, the class-specific proportion factor is compared over time, multiplied by an appropriate class-confined house price index. This class-confined house price index is different from the overall house price index. Consequently it is not recommended to simply use the overall house price index to measure the price change in the various classes.

The price index for a fee with multiple price-dependent rates can be compiled knowing only the distribution of the prices in the base period. More particularly, index formulas have been provided assuming a log-normal distribution of the house prices. These index formulas are rather easy to apply in practice if no micro-data on prices is available or if the use of micro-data is not satisfactory. Moreover, it makes the index more transparent for users. This might help communicating about the results as the price index for the fee can behave significantly different from the house price index even if the fee scheme itself is kept constant. Although care should be taken while estimating the initial price distribution, in our example the final results are quite robust with respect to small changes made on the parameters of the distribution.

In practice, if the log-normal assumption seems to be significantly put in doubt, then the index formulas developed here are not valid anymore. If necessary, our approach can be generalized to cope with alternative distributions. In any case, it is crucial to acknowledge that the distribution of house prices plays a major role in the compilation of a price index for fees with class-specific rates.

References

[1] COMMISSION (2001): Commission regulation (EC) No 1920/2001 of 9 December 1998.

[2] COMMISSION (1998): Council regulation (EC) No 2166/1999 of 8 October 1999.

[3] RIBE M. (1999): Treatment of Income-Dependent Prices in a CPI, Proceedings of the Measurement of Inflation Conference, Cardiff, pp.434-440.

[4] DALÉN J. (1995): Swedish CPI practices for difficult areas with a note on CPI Subindexes for services with income-dependent fees, 2nd meeting of the International Working Group of Price Indices, Stockholm.

[5] MACKAY N. (2011): London house prices are power-law distributed, Department of Mathematics, University of York, http://arxiv.org/abs/1012.3039v2.

[6] OHNISHI T., MIZUNO T., SHIMIZU C., WATANABE T. (2010): On the evolution of the house price distribution. Understanding Inflation Dynamics of the Japanese Economy, working paper series no. 56.

[7] EUROSTAT (2011): Technical manual on owner-occupied housing for Harmonised Index of Consumer Prices.

Appendix 1

Let P_i^0 and P_i^t be two random variables. First, we define two new random variables as follows:

$$\hat{P}_t^i = \begin{cases} P_t^i & \text{if } l_t^s \leq P_t^i < u_t^s \\ 0 & \text{otherwise} \end{cases} \quad \hat{P}_0^i = \begin{cases} P_0^i & \text{if } l_0^s \leq P_0^i < u_0^s \\ 0 & \text{otherwise} \end{cases}$$

We suppose that there are in total n prices. The total number of transactions is fixed and is not a random variable.

$$\begin{aligned}
HPI_t^s &\equiv E\left[\frac{\sum_{l_t^s \leq P_t^i < u_t^s} P_t^i}{\sum_{l_0^s \leq P_0^i < u_0^s} P_0^i}\right] \\
&= E\left[\frac{\sum_{i=1}^n \hat{P}_t^i}{\sum_{i=1}^n \hat{P}_0^i}\right] \\
&\approx \frac{E[\sum_{i=1}^n \hat{P}_t^i]}{E[\sum_{i=1}^n \hat{P}_0^i]} \\
&= \frac{n \cdot E[\hat{P}_t^i]}{n \cdot E[\hat{P}_0^i]} \\
&= \frac{E[\hat{P}_t^i]}{E[\hat{P}_0^i]} \\
&= \frac{E[\hat{P}_t^i | l_t^s \leq P_t^i < u_t^s] \cdot P(\hat{P}_t^i | l_t^s \leq P_t^i < u_t^s)}{E[\hat{P}_0^i | l_0^s \leq P_0^i < u_0^s] \cdot P(\hat{P}_0^i | l_0^s \leq P_0^i < u_0^s)} \\
&\quad \frac{+E[\hat{P}_t^i | P_t^i < l_t^s \text{ or } P_t^i \geq u_t^s] \cdot P(\hat{P}_t^i | P_t^i < l_t^s \text{ or } P_t^i \geq u_t^s)}{+E[\hat{P}_0^i | P_0^i < l_0^s \text{ or } P_0^i \geq u_0^s] \cdot P(\hat{P}_0^i | P_0^i < l_0^s \text{ or } P_0^i \geq u_0^s)} \\
&= \frac{E[P_t^i | l_t^s \leq P_t^i < u_t^s] \cdot P(P_t^i | l_t^s \leq P_t^i < u_t^s)}{E[P_0^i | l_0^s \leq P_0^i < u_0^s] \cdot P(P_0^i | l_0^s \leq P_0^i < u_0^s)}
\end{aligned}$$

Appendix 2

If $f(x)$ is a probability density function (pdf), a partial expectation is defined as follows:

$$E[X|X \geq k] \cdot P(X|X \geq k) = \int_k^{+\infty} xf(x)dx.$$

In the case of a log-normal distribution with parameters m and s, the pdf has the following form:

$$f(x) = \frac{1}{xs\sqrt{2\pi}} \exp\left(-\frac{(ln(x)-m)^2}{2s^s}\right).$$

Consequently, we have here:

$$E[X|X \geq k] \cdot P(X|X \geq k) = \int_k^{+\infty} \frac{1}{s\sqrt{2\pi}} \exp\left(-\frac{(ln(x)-m)^2}{2s^s}\right) dx.$$

We now apply a transformation on the variable x:

$$y = \frac{ln(x)-m}{s} \qquad dx = s\exp(sy+m)dy$$

This leads to the following integral:

$$\begin{aligned}
& \int_{y=\frac{ln(k)-m}{s}}^{+\infty} \frac{1}{s\sqrt{2\pi}} \exp\left(-\frac{1}{2}y^2\right) s\exp(sy+m)dy \\
= & \int_{y=\frac{ln(k)-m}{s}}^{+\infty} \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{1}{2}y^2 + sy + m\right) dy \\
= & \int_{y=\frac{ln(k)-m}{s}}^{+\infty} \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{1}{2}(y^2 - 2ys + s^2) + (m + \frac{1}{2}s^2)\right) dy \\
= & \frac{1}{\sqrt{2\pi}} \exp\left(m + \frac{1}{2}s^2\right) \int_{y=\frac{ln(k)-m}{s}}^{+\infty} \exp\left(-\frac{1}{2}(y-s)^2\right) dy.
\end{aligned}$$

We apply another transformation on variable y:

$$v = y - s \qquad dy = dv$$

After this transformation, the integral looks as follows:

$$\exp\left(m + \frac{1}{2}s^2\right) \frac{1}{\sqrt{2\pi}} \int_{v=\frac{ln(k)-m}{s}-s}^{+\infty} \exp\left(-\frac{1}{2}v^2\right) dv.$$

The pdf of a standard normal distribution is $\frac{1}{\sqrt{2\pi}}\exp(-\frac{1}{2}x^2)$.Consequently, the last term can be expressed using the cumulative distribution function (cdf) of a standard normal distribution, denoted by $\phi()$:

$$\exp\left(m + \frac{1}{2}s^2\right)\left(1 - \phi\left(\frac{ln(k)-m}{s} - s\right)\right).$$

A known property is that $1 - \phi(x) = \phi(-x)$. This allows us to conclude that:

$$E[X|X \geq k] \cdot P(X|X \geq k) = \exp\left(m + \frac{1}{2}s^2\right) \phi\left(\frac{m + s^2 - ln(k)}{s}\right).$$

Appendix 3

The index is based on the assumption that the parameters of the log-normal distribution are $m_0 = 12.69$ and $s_0 = 0.44$. If the fee scheme is kept constant, the index is compiled as follows:

$$I_t = \widetilde{w}^1 \cdot \frac{f_t^1}{f_0^1} \cdot \frac{\widetilde{E}_t^1}{\widetilde{E}_0^1} + w^2 \cdot \frac{r_t^2}{r_0^2} \cdot \frac{E_t^2}{E_0^2} \cdot HPI_t$$

where:

- $\widetilde{E}_t^1 = \phi\left(\frac{ln(287143) - m_0 - ln(HPI_t)}{s_0}\right), \forall t = 0, 1, \ldots$
- $E_t^2 = \phi\left(\frac{m_0 + s_0^2 - ln(287143) + ln(HPI_t)}{s_0}\right), \forall t = 0, 1, \ldots$
- $f_t^1 = 100, \forall t = 0, 1, \ldots$
- $r_t^2 = 7\% - 20000 \cdot \frac{1 - \widetilde{E}_t^1}{E_t^2 \cdot e^{m_0 + 0.5 s_0^2 + ln(HPI_t)}}, \forall t = 0, 1, \ldots$
- $\widetilde{w}^1 = \frac{f_0^1 \cdot \widetilde{E}_0^1}{f_0^1 \cdot \widetilde{E}_0^1 + r_0^2 \cdot E_0^2 \cdot e^{m_0 + 0.5 s_0^2}}$
- $w^2 = \frac{r_0^2 \cdot E_0^2 \cdot e^{m_0 + 0.5 s_0^2}}{f_0^1 \cdot \widetilde{E}_0^1 + r_0^2 \cdot E_0^2 \cdot e^{m_0 + 0.5 s_0^2}}$