CAHIER D'ÉTUDES
WORKING PAPER

N° 160

OBJECTIFIED HOUSING SALES AND RENT PRICES IN REPRESENTATIVE HOUSEHOLD SURVEYS: THE IMPACT ON MACROECONOMIC STATISTICS

M. DENISA NAIDIN SOFIE R. WALTL
MICHAEL H. ZIEGELMEYER

JUNE 2022

BANQUE CENTRALE DU LUXEMBOURG
EUROSYSTÈME

Objectified Housing Sales and Rent Prices in Representative Household Surveys: the Impact on Macroeconomic Statistics

M. Denisa Naidin[a,b], Sofie R. Waltl*[a,c], and Michael H. Ziegelmeyer[d,e]

[a] *Luxembourg Institute of Socio-Economic Research (LISER)*

[b] *University of Luxembourg*
[c] *Vienna University of Economics and Business*
[d] *Banque centrale du Luxembourg (BCL)*
[e] *Munich Center for the Economics of Ageing (MEA)*

Abstract

Reliable macroeconomic housing and wealth statistics as well as counterfactual analyses across housing tenure status require hypothetical sales and rent prices for properties off the market reflecting current market conditions and representing the entire housing stock. We replace subjective values reported by participants in the *Luxembourg Household Finance and Consumption Survey* by objectified values imputed via hedonic models estimated on observable market data. We find that the participants' tendency to over- and under-report values is strongly correlated with tenure length, tenure type, dwelling type, household income and wealth. We find shifts in the wealth distribution, detect large regional variation in price-to-rent, price-to-income and rent-to-income ratios as well as significant affordability concerns: only 18% of all renting households could theoretically afford to purchase the dwelling they currently rent at market conditions. These renters are usually younger, at the top of the wealth and income distribution, and reside outside Luxembourg City.

Keywords: Macroeconomic Statistics; Subjective Assessments; Surveys; Measurement Errors; Housing and Rent Markets; Housing Wealth; Affordability
JEL codes: E58; G51; R21; R31.

Notes and Acknowledgements: This research benefits from funding by the *FNR Luxembourg National Research Fund*, CORE Grant No. 3886 (ASSESS). We are grateful for fruitful discussions and comments by participants of presentations at the *National Bank of Slovakia*, the *LISER GeoFin PhD Meeting*, the *Banque centrale du Luxembourg*, the *2021 ESCoE Conference on Economic Measurement*, the *European Central Bank (DG Statistics)*, the *2021 Research Seminar of the Household Finance and Consumption Network*, the meeting of the *Expert Group on Distributional Financial Accounts* managed by the *European Central Bank*, the *2022 ASSA Annual Virtual Meeting*, the *Luxembourg Institute of Socio-Economic Research (LISER)* and the 2022 Austrian and Slovenian Statistical Days at the *Graz University of Technology.* We are grateful for valuable information and additional data provided by Julien Licheron (LISER), François Peltier and Joe Peiffer (STATEC), and Myriam Wiltzius (Luxembourg Ministry of Finance). We would like to thank Gaston Reinesch, Paolo Guarda, Bob Kaempff and the members of the Comité de relecture (all BCL) for useful comments.
This article uses data from the *Luxembourg Observatoire de l'Habitat* and the *Luxembourg Household Finance and Consumption Survey* (3rd wave).
This article should not be reported as representing the views of the BCL or the Eurosystem. The views expressed are solely those of the authors.

*Corresponding author: Sofie R. Waltl; Maison des Sciences Humaines; 11, Porte des Sciences; 4366 Esch-sur-Alzette/Belval; Luxembourg; sofie.waltl@liser.lu; www.sofiewaltl.com

Non-technical Summary

In Luxembourg, as elsewhere, real estate is generally the most important asset held by households. Determining total wealth, as well as its distribution across individuals, households, groups of households and countries, requires accurate estimates of the current value of housing assets. In addition, macroeconomic housing statistics need reliable estimates of the housing stock reflecting current market conditions. However, at any given time only a small share of the housing stock is available for sale or rent. A proper macro-economic perspective should consider how evolving market conditions affect the value of the entire stock of residential real estate.

Wealth surveys are often used to close this gap, but they rely on homeowners and renters guessing the current sales or rent price of their off-market dwellings. Given detailed information on individual households, wealth surveys can provide insightful housing statistics as long as owners and renters accurately report current market values. However, previous studies have shown that owner-reported values provided in surveys track changes well, but misreport levels.

The *Luxembourg Housing Observatory* collects comprehensive data on advertised dwellings for sale and rent in Luxembourg. We conduct a technical analysis, using housing characteristics to match market data with responses from the *2018 Luxembourg Household Finance and Consumption Survey (LU-HFCS)*. Advertisements provide detailed data on dwelling characteristics and reflect current market conditions and supply-side sentiment. To improve the match, we extended the LU-HFCS with questions on detailed physical and locational dwelling characteristics. Regardless of tenure status, we then estimate "objectified" current market sales and rent prices for each dwelling in the LU-HFCS using hedonic imputation models estimated on the advertisement data of the Housing Observatory, which we compare to the responses provided in the survey.

We use the weighting scheme included in the LU-HFCS to gross up the results to macroeconomic statistics. This exercise leads to four main results:

- First, the median value of net wealth reported by Luxembourg's owner-occupiers is about EUR 50,000 lower than an estimate based on "objectified" values of their residential real estate. The differences between reported and imputed values are substantial for households in the lowest net income quintile but not significant in the top quintile.
- Second, we characterise survey participants who tend to under- or over-report the value of their home. The reliability of the self-reported value depends on knowledge about the local housing market and their ability to adapt these general trends to the attributes of their own home. We find strong correlations with tenure length, tenure type, type of dwelling, household income and wealth.
- Third, we calculate dwelling-level price-to-rent ratios, price-to-income ratios and rent-to-income ratios to then obtain representative values for the entire housing stock. We find a substantial variation across households and regions.
- Fourth, we perform a micro-simulation to assess whether renters could theoretically purchase their current dwelling. This exercise takes account of renters' financial and living conditions, as well as the institutional framework and housing market conditions. On the basis of imputed sales prices, we conclude that only 18% of these households could repay the mortgage required to purchase the dwelling they currently rent.

Résumé non-technique

Au Luxembourg, comme ailleurs, les ménages détiennent la plupart de leur patrimoine sous la forme d'actifs immobiliers. Par conséquent, des estimations fiables de la valeur courante des actifs immobiliers sont indispensables pour déterminer tant le patrimoine global des ménages que sa répartition à travers les ménages individuels, différentes catégories de ménages ou même différents pays. De plus, ces estimations doivent couvrir l'ensemble du parc immobilier, pas uniquement la partie qui est disponible à la vente ou à la location à un moment déterminé. Ainsi, une perspective macro-économique appropriée devrait prendre en compte comment l'évolution du marché affecte la valeur de l'ensemble du patrimoine immobilier résidentiel.

Pour combler ces limitations des statistiques macroéconomiques, les chercheurs ont souvent recours aux enquêtes sur le patrimoine des ménages, lors desquelles certains propriétaires ou locataires peuvent fournir une indication subjective de la valeur courante des immeubles qu'ils habitent. Ces enquêtes collectent également des informations détaillées sur les ménages interrogés, ce qui pourrait permettre d'apporter plus de précision aux valeurs immobiliers qu'ils rapportent lors de l'enquête. En effet, des études antérieures ont montré que les valeurs rapportées par les propriétaires évoluent avec les variations des prix des biens immobiliers, mais divergent souvent du niveau de ces prix.

Au Luxembourg, *l'Observatoire de l'Habitat* collecte des données sur les logements à partir des annonces pour la vente ou la location. Dans cette analyse technique, nous combinons ces données avec celles issues de la dernière Enquête sur le comportement financier et de consommation des ménages (LU-HFCS) conduite en 2018. Les annonces immobilières fournissent des données détaillées sur les logements et reflètent les conditions du marché et le sentiment du côté de l'offre. L'édition 2018 de l'enquête LU-HFCS posait des questions additionnelles concernant les caractéristiques physiques et la localisation des logements pour mieux établir un lien avec les données de l'Observatoire. Nous estimons des modèles hédoniques avec les données de l'Observatoire pour ensuite imputer la valeur "objectivée" de vente ou de location pour chaque logement couvert par l'enquête LU-HFCS, que nous comparons ensuite aux réponses fournies lors de l'enquête.

En appliquant les pondérations de l'enquête LU-HFCS, nous pouvons extrapoler de l'échantillon représentatif qui a participé à l'enquête à l'ensemble de la population résidente au Luxembourg. Cet exercice conduit à quatre résultats principaux :

- Premièrement, les propriétaires-occupants au Luxembourg tendent à sous-déclarer la valeur de leur patrimoine immobilier par rapport aux valeurs "objectivées" (différence de 50,000 euros pour le ménage moyen). Les différences entre les valeurs déclarées et imputées sont substantielles pour les 20% des ménages avec des revenus plus modestes, mais la différence n'est pas significative pour les 20% des ménages les plus aisés.
- Deuxièmement, nous nous focalisons sur les participants à l'enquête qui ont tendance à dévier par rapport à la valeur "objectivée" de leur logement. La fiabilité de la valeur auto-déclarée dépend de la connaissance du marché local du logement et de la capacité à adapter ces tendances générales aux attributs du propre logement. Les déviations positives ou négatives sont fortement corrélées avec la durée d'occupation, le type d'occupation, le type de logement, le revenu et le patrimoine du ménage.
- Troisièmement, pour chaque logement nous calculons le ratio prix/loyer, le ratio prix/revenu et le ratio loyer/revenu. Ces ratios varient fortement à travers les ménages et les régions du pays.
- Quatrièmement, nous conduisons une micro-simulation pour évaluer si les locataires pourraient théoriquement acheter le logement qu'ils occupent actuellement. Cet exercice tient

compte de la situation financière des ménages locataires, ainsi que du cadre institutionnel et des conditions du marché immobilier. Sur base des prix objectivés, nous trouvons que seulement 18% de ces ménages seraient capables de rembourser l'hypothèque nécessaire pour l'achat de la résidence qu'ils louent actuellement.

1 Introduction

At any given moment, the vast majority of dwellings forming a country's housing stock is neither on sale nor available for rent. Thus, these dwellings' market value is uncertain since they have not recently undergone a market matching process.[1]

While this is also true for other asset classes, the enormous amount of private wealth tied up in real estate[2] requires particularly accurate estimates for this asset class. *Every* dwelling is unique given its combined set of physical and locational characteristics, and, thus, real estate needs to be treated carefully: prices of other "comparable" homes recently sold or rented only give an indication but no definite guideline for a market value of a dwelling currently off the market. This is especially relevant for markets with a heterogeneous housing stock. In Luxembourg, the focus of our study, more than 40% of the housing stock is at least half a century old.[3]

In addition, a housing unit cannot simultaneously be active on the rent and sales market.[4] Thus, at least either the sales price or the rent remains unknown. This naturally occurring fact, however, limits the compilation of macroeconomic housing statistics targeting the entire stock and reflecting current market conditions.

In our technical analysis, we fill this gap by combining a country-representative survey with actual market data to compile objectified macroeconomic statistics. Sophisticated survey weighting schemes allow us to gross up results to population totals and thus also price statistics for the owner-occupied and rented housing stock. Thus, we introduce a new policy tool for simulating consequences of various housing market scenarios.

We focus on the *Luxembourgish (LU-HFCS)* part of the *Household Finance and Consumption Survey* harmonising core variables across many European countries.[5] Survey data have been widely used before for similar purposes. For instance, Garner and Verbrugge (2009) make use of the ample information within U.S. household surveys to compile macroeconomic housing statistics. Other studies use surveys to compile wealth statistics including housing wealth components (see, e.g., EG-LMM, 2020). However, these studies are limited to self-reported house values and do not use any micro-level market-data except for summary plausibility checks.

This appears problematic as values collected in wealth surveys are known to be prone to several sources of imprecision (see, for instance, Vermeulen, 2018; Waltl and Chakraborty, 2022; Bach et al., 2019; Kennickell, 2019; Waltl, 2022). After reviewing studies focusing on values reported by owners, Agarwal (2007) concludes: "there is general agreement [...] that homeowners significantly misestimate their house value". He reports substantial average absolute mis-estimation (mainly focusing on the U.S.) ranging between 14% and 25% (see Kish and Lansing, 1954; Kain and Quigley, 1972; Goodman Jr and Ittner, 1992; Kiel and Zabel, 1999; Agarwal, 2007; Benítez-Silva et al., 2015; Molloy and Nielsen, 2018; Lepinteur and Waltl, 2021, for detailed results). However, price changes measured by subjective and objective house price indices seem to follow similar dynamics (see Kiel and Zabel, 1997; Mathä et al., 2017). Lepinteur and Waltl (2021) find that housing market trends are well tracked when conducting systematic convergent validity tests on pooled estimates of all participants in national surveys in the U.S. and Europe. In contrast, the level of estimates appears to be systematically biased.

The goal of this article is thus to perform a micro-level assessment of the reliability of owner or renter reported current market sales and rent prices. We characterise the type of deviation and how they are linked with characteristics of the surveyed household. Further, we measure the magnitudes of deviation, and demonstrate the potential of using objectified sales prices and rents to construct macroeconomic statistics including counterfactuals when hypothetically switching the tenure type.

For this purpose, we use a rich set of market sales and rent prices collected by the *Luxembourg Housing Observatory*.[6] We match the housing characteristics available for market data with those reported in the survey. To enable a better match, we extended the Luxembourg HFCS survey conducted in 2018. We then estimate hedonic valuation models on market data, match input variables to survey questions and evaluate these models for all household main residences in the LU-HFCS, thus obtaining *objectified* current sales and rent prices. Similar imputations are occasionally performed for missing or counterfactual rents for owner-occupied dwellings in the *United States' Panel Study of Income Dynamics (PSID)*, the *Household, Income and Labour Dynamics in Australia (HILDA)* and the *German Socio-Economic Panel (SOEP)* as highlighted in Alexeev (2020).

We then use these imputed values to compile several macroeconomic statistics: total (housing) wealth, distributional break-downs, aggregate and regional price-to-rent ratios, price-to-income ratios and rent-to-income ratios, as well as housing affordability measures. We find that on average imputed prices are slightly higher than those reported by homeowners. However, renters strongly under-report the sales potential of their currently rented unit. These rather small deviations observable for single units translate into substantial increases in aggregate effects and, even more importantly, changes in the distribution: median net worth of owner-occupiers increases by almost EUR 50,000; however, in the lowest income quintile the increase amounts to EUR 170,302, in the second-highest income quintile to EUR 20,448 and even a small decrease for the top 20%. Similar, large changes are also observed along the wealth distribution. Since the LU-HFCS is part of the European *Household Finance and Consumption Survey*, we perform simulations for several countries. Data for countries with large shares of owner-occupiers and expensive housing stocks may be prone to substantial "corrections" in net wealth measures.

Our imputation strategy also allows us to assess the secondary housing market in terms of affordability by running a counterfactual study. The results suggest that current market conditions would make it almost impossible for a large fraction of renters to purchase the home they currently rent using their wealth and income. This finding is consistent with the recent blog by BCL Governor Gaston Reinesch (2022). For 50%, such a purchase would require financial resources equal to twelve total annual net household incomes (excluding any transaction and financing costs). Only 15% of all renting households could theoretically afford the purchase and would also economically benefit from doing so. Roughly 17.5% would have to pay less interest on a hypothetical mortgage than their current rent. These households thus could greatly benefit from home-ownership but are hampered by lacking wealth to overcome the down-payment constraint. If house prices rise faster than rents, the ratio of renters to owners should increase over time.

Similar to the simulations we perform here, the established toolkit linking survey observations to real estate market data allows policy-makers and researchers in general to micro-simulate effects of policies (e.g., targeted to support either home-ownership or renting), stress-test households' portfolios in hypothetical scenarios, or simply bringing housing statistics forward or backwards in time. For the latter, the housing stock described in the survey would be kept constant (as the housing stock anyhow is fixed in the short-run) but sales and rent prices could flexibly evolve in-line with changing market conditions by adjusting the time window the market data is retrieved from. By that, Luxembourg's current set of population-representative housing statistics would be vastly extended.

The remainder of this article is structured as follows: section 2 presents the data sources used, section 3 describes the imputation strategy applied, and section 4 demonstrates impacts on selected macroeconomic statistics. Finally, section 5 concludes. A comprehensive appendix provides additional details and the attached online appendix supplemental materials.

2 Data Sources

2.1 The LU-HFCS

The LU-HFCS is part of the Pan-European HFCS initiative conducting ex-ante harmonised wealth surveys. Our analyses is based on 1,616 households participating in the 3rd wave of the HFCS conducted in Luxembourg in 2018 (see the technical report by Chen et al., 2020, for further details). Answers of interviewed households are weighted such that results are representative for the population of households residing in Luxembourg.[7] Table O.2 in the appendix reports summary statistics.

To achieve a comprehensive wealth measure, the HFCS always asks owner-occupiers to estimate their home's current market price. Renters, in contrast, are asked for the currently paid monthly rent. While the first reflect subjective beliefs, the latter might at most be affected by reporting errors, not subjectivity.[8] It is important to note that survey participants are, in general, no experts when it comes to real estate markets. Albeit people interested in buying or selling a home are likely to gather extra information by observing the market and getting expert opinions by a real estate professional,[9] such costly endeavours are very unlikely to be undergone when preparing for a general household survey interview. Thus, responses arguably reflect beliefs and potentially also wishes.

Exclusively the 3rd wave of the LU-HFCS contains questions about hypothetical counterfactual prices, as well as an additional set of questions eliciting physical and locational characteristics of a household's main residence. On top, survey interviewers add supplemental assessments of the dwelling and surroundings. Regarding the former, owner-occupiers are asked for a hypothetical monthly rent and all respondents are tasked to estimate a hypothetical current sales price. We print the exact phrasing of the questions eliciting these values in Appendix O.2. Regarding the latter, the following physical characteristics of the main residence are retrieved during the survey interview usually taking place at the participant's home thus minimising stark reporting errors: surface, type of residence, plot size, year of construction, number of bedrooms and the energy class. The survey interviewer provides an additional assessment of the structure and neighbourhood.

In terms of location, we included a question asking for the postcode during the interview. Luxembourg has a strikingly detailed system of postcodes: While small municipalities in the countryside share one code, in Luxembourg City (almost) every street has a separate code. Longer streets are even split into several codes. Overall, there are 4,022 regular postcodes in a country of 2,586 km^2 and roughly 600,000 residents.[10]

This additional information allows us to impute realistic market values based on a hedonic valuation model for every main residence appearing in the survey. We do so for both, owner-occupiers (represent 69% of the households in the sample or 1,207 unweighted observations), and renters or households using their residence free of charge (represent 31% of the households in the sample or 409 unweighted observations).

Personal characteristics refer to the financially most knowledgeable person in the household who acts as main interview partner (reference person). The LU-HFCS data set is multiply imputed for all variables. Reported point estimates are the average of the weighted point estimate across five implicates. The variance of our estimators is estimated using standard bootstrap variance based on a set of 1,000 replicate weights adjusted for the between and within imputation variance of the five multiply imputed implicates. The estimation uncertainty of our hedonic valuation models to impute objectified current sales and rent prices is indirectly taken into account via bandwidths used for classifying accurate versus over- or under-reported values.

Few ambiguities regarding the exact location of dwellings are reflected via heterogeneity across implicates.

2.2 Market data

We make use of a comprehensive data pool consisting of advertised dwellings for sale and rent in Luxembourg. The data base is maintained by LISER to serve as *Housing Observatory*. Table O.3 in the appendix reports basic summary statistics.

We estimate our imputation models on advertised sales prices net of taxes and fees, and advertised rent prices net of utilities and other charges. As explanatory variables, we use all characteristics available for both, survey and market data. Regarding time frame, we take all adverts posted during the LU-HFCS fieldwork period; for comparison, we also consider all the notary deeds recorded for dwelling transactions during the same period (see Table 1).

Given the delay between price setting and official recording of transactions by notaries, advertisements provide a timely snapshot of current market conditions and seem hence to be a practical choice. Moreover, potential alternative data sources, such as sales deeds, contain very little information on dwelling characteristics.

Rent contracts are not officially recorded in Luxembourg. Thus, there is no external benchmark for advertised rents. This, however, seems less of an issue as, unlike for sales prices, bargaining rents is rather uncommon and an advertised rent usually matches the actual rent paid (see, for instance, Hill and Syed, 2016).

Nonetheless, extra checks are needed to justify using advertised sales prices due to potential price negotiations and, thus, advertised sales prices may be higher than what could realistically be achieved on the market. Kolbe et al. (2020) find for Berlin that asking prices retrieved from advertisements constitute an upper bound for the final sales price. They argue that listings with excessively high asking prices are often not successful and thus do not lead to a deal under the proposed conditions. Additionally, a common sales strategy is the following: an unrealistically high price is initially posted. If unsuccessful, the advertised price is step-by-step decreased until a buyer willing to pay this price is found. Furthermore, negotiations between potential buyers and the seller may lead to further decreases. Yet, in tight markets like the one in Luxembourg substantial price reductions are unlikely.

To validate our data source, we thus compare advertised sales prices to final prices for apartments recorded by notaries. We restrict the analysis to sales as there is no official rent registry and to apartments as house transactions are often split across several notary items (e.g., separate ones for the house itself and an attached garage, cellar and garden) unfortunately not uniquely identifiable to belong to the same property. The same is true for apartments (see below) yet they are usually less splits across lines in notary deeds than houses. Adverts and the HFCS, in general, refer to complete bundles.

Furthermore, to distinguish dwellings for own use (relevant for our study) from buy-to-let transactions, we requested additional information be retrieved from the notary act. This variable specifies whether buyers make use of the *Bëllegen Akt* providing a tax credit for owner-occupiers (see subsection 4.3 for details about this policy). This tax credit is available only once in a person's life-time and, thus, the resulting sub-sample relates best to the notion of first-time "home-buyers" rather than buy-to-let activities.

Table 1 compares advertised and recorded final sales prices per square meter and relates them to each other via (unpaired) Pearson correlation coefficients. As result, we obtain $\rho = 98.38\%$ and, thus, reassuringly document almost perfect co-movement of prices reported in the two

Table 1 – Adverts

	Adverts	Notary Deeds
Price per unit [EUR/m^2]		
Q1	4,774.81	4,227.05
Median	5,748.91	5,188.47
Mean	6,462.81	5,550.55
Q3	7,506.32	6,475.64
IQR	2,731.51	2,248.59
Std. Dev.	2,364.86	2,024.04
Dwelling surface [m^2]		
Q1	71.00	70.00
Median	87.16	84.66
Mean	91.65	88.46
Q3	106.00	101.35
IQR	35.00	31.35
Std. Dev.	31.63	29.67
Number of Observations	8,381	4,737

Notes: The table reports summary statistics for advertisements and notary deeds for characteristics reported in both sources. The data refers to flats advertised or sold (registered as for own use) in Luxembourg in 2018. *Q*1 and *Q*3 abbreviate quartiles.
Source: Luxembourg's Housing Observatory.

sources. The distributions of prices and dwelling surface recorded by notaries is, as a whole, shifted downwards as compared to advertisements. This shift is significant in statistical terms: According to a one-sided Mood's median test, the values reported in advertisements are statistically significantly higher at the 5% level. Furthermore, a *Mann-Whitney U Test* is significant at the 1% level indicating that the two populations are nonidentical level-wise.

While values in advertisements are higher than data reported in notary deeds, they still offer an overall accuracy comparable to notary deeds. First, the real estate market is very tight in Luxembourg so that usually the advertised price equals the transaction price and the listing time is short.[11] Second, notary deeds contain also transactions that have never been advertised because real estate is sold to relatives or friends based on favourable conditions and, thus, not describing true market prices. Third, notary deeds reflect the market with a delay of several months as a consequence of the lengthy gap between price agreement and final recording of the transaction. Recording with notaries only happens at the very end of the purchase process (which also requires a successful mortgage application). Fourth and most importantly, notary deeds do not always refer to complete bundles of transacted goods, i.e., a unit sold may be spread over more than one line in the notary deeds due to the merging of previously separated units at some point in the past or additional acquisition of garages, other types of storage facilities or outside space.

Adverts' information content, though containing more dwelling characteristics than notary deeds, is still not massive. Yet, the single most important information is available: location. Thus, we can make extensive use of geographic data as explained below. We have selected the municipal (commune) level (the finest administrative area available in Luxembourg) as our main unit of analysis due to the opportunities it offers to proxy local phenomena and policies, but also to connect into other data sets available at this scale. Nonetheless, Luxembourg's two main urban agglomerations – Luxembourg City and Esch-sur-Alzette – are significantly larger population-wise compared to any other municipality.[12] Given this disparity in population distributions, we break down the larger urban areas into smaller but administratively relevant neighbourhood units. Precisely, Luxembourg City is divided into its composing neighbourhoods and Esch-sur-Alzette is separated from its largest – and truly distinct – suburb Belval. For all other locations we preserve the municipal level of analysis.

We label the result "composite geography" which we subsequently match to postcodes (the unit of collection in the LU-HFCS) via GIS juxtaposition between postcode centroids and the boundaries of municipal/neighbourhood units. As Luxembourgish postcodes occasionally cross administrative boundaries, the centroids are adapted to cover administrative sections of postcodes. The resulting conversion table can therefore assign multiple municipalities or neighbourhoods to a single code. To recognise this variance in the municipal adherence of an HFCS observation, we introduce for these few observations locationally varying implicates.

This location information ultimately enters the hedonic models in three different ways: dummies, a distance measure and a non-parametric spline. The first option estimates for each locational entity a separate shadow price. The second option follows Glaesener and Caruso (2015) in calculating the approximate travel distance from each dwelling to Luxembourg City. Finally, we estimate a locational spline as a *Markov Random Field Smooth* (see Wood, 2006) that links each geographical entity part of the composite geography to all of its neighbours thus measuring potential spatial lags. As a consequence, the model estimation technique switches in the latter case from standard maximum likelihood to penalised maximum likelihood.

2.3 Comparing Market and Survey Data

Table O.2 describes dwelling characteristics as reported in the HFCS data set and Table O.3 reports summary statistics for advertised units. Summary statistics for the survey data are population-representative and thus a comparison reveals differences between the current stock of housing and the mix of dwellings currently on the market. One key insight from this comparison is that properties currently on the market tend to be considerably smaller and are less likely to come with additional private outside space (e.g., a garden). For both data sets, the mean and median age of main residences of renters and owner-occupiers are almost identical.

Looking at differences between dwellings associated with tenure status, we find overall similar deviations: dwellings for sale or already owner-occupied are on average substantially larger than those rented. While rented properties in the two sets are comparable in size, properties owned/sold are noticeably smaller in the pool of advertised properties than among the residences of HFCS respondents. This translates not only into more bedrooms per property, but also into larger plot sizes.

3 Imputing Market Prices

3.1 Strategy

Based on the market data presented in subsection 2.2, we estimate hedonic imputation models by regressing prices on price-determining characteristics (Rosen, 1974; Rosen et al., 1984). Here, we briefly describe the applied concepts and describe our model selection procedure taking into account various aspects relevant for imputing hypothetical prices using this model. Full methodological details are provided in Appendix A.

We make use of information available for both survey and market data. To accommodate differences in recording and coding style, we present a correspondence table in the appendix (Table 13). Most information is provided by survey respondents, yet the HFCS interviewer conducting the interview at the household's main residence adds supplemental assessments of the overall status of the dwelling's structure and surrounding area.

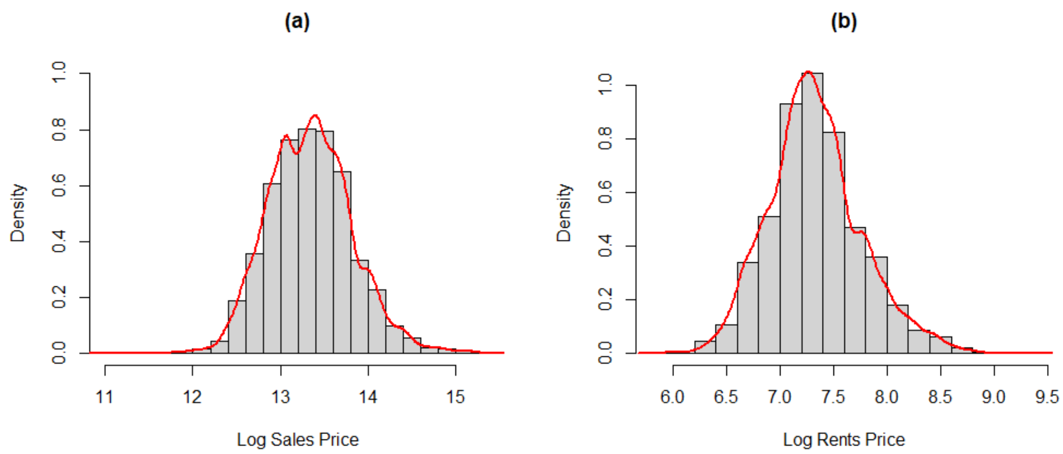
The general model regresses logged sales $P^S_{ht} \in \mathbb{R}^n_s$ or logged monthly rent $P^R_{ht} \in \mathbb{R}^n_r$ prices observed at time t for dwellings $h = 1, \ldots, n$ on a set of corresponding physical and locational characteristics:

$$\log P^{S \vee R}_h = \beta X_h + \lambda L_h + \varepsilon_h, \tag{1}$$

where $n \in \{n_S, n_R\}$ and $n_S, n_R \in \mathbb{N}$ denotes the number of sales or rent observations, respectively, $X \in \mathbb{R}^{n \times k}$ and $L \in \mathbb{R}^{n \times l}$ matrices containing $k, l \in \mathbb{N}$ physical and locational dwelling characteristics with associated parameters $\beta' \in \mathbb{R}^k$ and $\lambda' \in \mathbb{R}^l$. Locational characteristics are incorporated as dummies, splines and distances, respectively. Details are provided in Appendix A. Finally, $\varepsilon \sim^{iid} \mathcal{N}(0, \sigma^2)$ denotes a normally and independently distributed error term with standard deviation $\sigma > 0$. The log-specification ensures an unbiased estimation of median sales and rent prices conditional on hedonic controls. Opting for a median-estimator accounts for the skewness inherent in both sales and rent prices (see Waltl, 2016). Our data confirms the appropriateness of this log-specification (see Figure 1). Predictions from model (1) are thus median prices for dwellings given observed price-determining attributes.

As alternative, a quantile regression specification would also yield median-unbiased predictions. However, including locationa as a spline linking neighbourhoods as done here is not yet developed for quantile models. As predictions from a quantile regression specification are very similar, we thus report them only in the appendix as a robustness check (see subsection A.4).

Figure 1 – Densities of Market Sales and Rent Prices.



Notes: The figure depicts empirical densities for market data of (a) logged sales and (b) logged rent prices. Silverman's rule of thumb is used for optimally selecting the bandwidth for smoothing the density line (see Silverman, 1986).

We estimate two sets (one for sales and one for rents) of each five specifications by step-wise varying sets of variables (Table 2). In particular, we consider three alternative specifications to account for spatial effects and test whether intuitively meaningful interaction effects are supported by the data (dwelling type × surface, as well as construction period × energy class). Table 14 in the appendix reports full numeric estimation results.

To check the models' performances we assess four dimensions: First, we plot imputed versus observed market prices for sales and rent prices. Reassuringly, Figure 2 suggests large correlations. Standard Pearson correlation coefficients are large and positive indicating co-movement and high in magnitude for both sales (87.19%) and rent (85.83%) prices.

Table 2 – Hedonic Model Specifications

	Main Model	Alt. 1	Alt. 2	Alt. 3	Alt. 4
Physical characteristics	✓	✓	✓	✓	✓
Interactions	✓	✓	✓	✗	✗
Locality Dummies	✓	✗	✗	✓	✗
Distance to Capital	✗	✓	✗	✗	✗
Linked Neighbour Spline	✗	✗	✓	✗	✗

Notes: Physical characteristics comprise all match-able characteristics specified in Table 13 except location. Interactions between dwelling type and surface as well as between construction period and energy class are included in three of the models. Locality means dummy variables hinting towards the composite geography obtained from combining municipalities with the neighbourhood level for the two largest urban areas in the country. The distance to the capital is retrieved from Glaesener and Caruso (2015). The linked neighbour spline employs smooth terms linking neighbouring entities of the composite geography.

Second, we assess a variety of goodness-of-fit measures reported in Table 3. For sales data, the main model clearly outperforms the four alternative specifications, no matter which measure we rely on.

Table 3 – Goodness-of-fit Assessment

	Adjusted R^2	AIC	BIC
Sales models.			
S.Main	0.824	-4,962.48	-3,783.79
S.Alt.1	0.703	2,282.41	2,509.08
S.Alt.2	0.819	-4703.96	-4039.10
S.Alt.3	0.809	-3,820.68	-2,725.11
S.Alt.4	0.507	9,414.07	9,550.08
Rent models.			
R.Main	0.744	-1,219.10	-164.11
R.Alt.1	0.679	307.96	514.81
R.Alt.2	0.728	-886.18	-624.54
R.Alt.3	0.742	-1,153.16	-174.02
R.Alt.4	0.526	3,160.34	3,284.46

Notes: The table reports goodness-of-fit measures for all hedonic models considered. AIC means the Akaike information criterion and BIC the Bayesian information criterion.
Source: Authors' calculations based on data from the *Housing Observatory*, on advertisements available between 1 January and 31 December 2018.

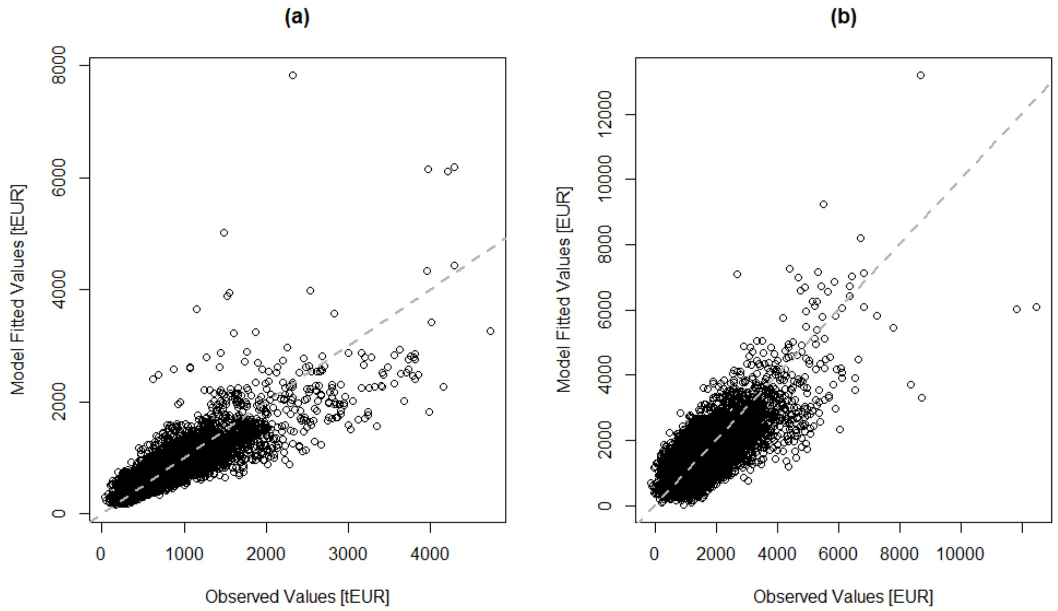
Third, we assess the models' predictive power by comparing overall *median/mean absolute errors (MedRE/MAE)* and the *median/mean relative errors (MedAE/MAE)* for sales S or rent R prices:

$$MAE(S \vee R) = \frac{1}{n}\sum_{i=1}^{n}\left|\hat{P}_i^{S\vee R} - P_i^{S\vee R}\right| \text{ and } MRE(S \vee R) = \frac{1}{n}\sum_{i=1}^{n}\left|\frac{\hat{P}_i^{S\vee R}}{P_i^{S\vee R}} - 1\right|,$$

$$MedAE(S \vee R) = \text{Median}\left|\hat{P}_i^{S\vee R} - P_i^{S\vee R}\right| \text{ and } MedRE(S \vee R) = \text{Median}\left|\frac{\hat{P}_i^{S\vee R}}{P_i^{S\vee R}} - 1\right|. \quad (2)$$

We assess the hedonic models' predictive power following a *out-of-sample* and *within-sample* procedure: for the first, we re-estimate the hedonic models leaving out 20% (i.e., 4,284) of all observations based on geographically-stratified random sampling and use the resulting reduced

Figure 2 – Correlation between imputed and observed prices.



Notes: The figure depicts observed and imputed market data for (a) sales and (b) rents prices using the models (S.Main) and (R.Main).
Source: Authors' calculations based on data from the *Housing Observatory*, on advertisements available between 1 January and 31 December 2018.

models to predict prices $\hat{P}^S$ and rents $\hat{P}^R$ via plug-in estimators. We compare the predicted values for these 4,284 units using the full and restricted model, respectively, and assess the ratio between the two. MAE ratios close to one and MRE ratios close to zero suggest that the model works well also for bundles of characteristics not necessarily found in the original estimation sample. This is important as we use the model for predicting prices for dwellings *per construction* off the market and thus not part of the estimation sample.

Table 4 reports the results: Out-of-sample and within-sample MAEs are comparable in size, reflected in ratios close to one. By construction, out-of-sample errors are expected to be larger than within-sample errors. This is indeed the case reflected by the ratio of out-of-sample to within-sample errors being consistently above – but reassuringly very close to – one.

The mean relative prediction error for dwellings excluded from the estimation sample amounts to roughly 15%. Hence, we must expect an error margin of this size when imputing sales and rent prices for survey observations. We also use this margin to identify over- or under-reported values as detailed in subsection 3.2.

Finally, we examine residuals by the most important price-determining characteristic: location. This checks whether imputations for certain locations need to be treated with extra caution. Figure 3 shows residuals per canton retrieved from the main models (S.Main and R.Main). A more detailed dis-aggregation is available in Figure O.1 in Figure A.4.

The width of interquartile ranges is quite similar across cantons and consistently overlaps zero. This holds true even when assessing the much more fine-grained split by our composite geography reflecting roughly municipalities (see Appendix A for a description of our applied spatial strategy).

We perform a robustness check by estimating the main specifications as quantile regression models (see subsection A.4 for details). The coefficients are very similar, so we use our hedonic linear model for the remaining analyses.

Table 4 – Predictive Power

	Out-of-sample (OoS) MAE (MedAE)	Out-of-sample (OoS) MRE (MedRE)	Within-sample (WS) MAE (MedAE)	Within-sample (WS) MRE (MedRE)	Ratio of absolute errors	Difference of relative errors
S.Main	109,688.7	0.1483	108,038.6	0.1465	1.0152	0.0018
	(61,288.5)	(0.1063)	(61,256.5)	(0.1055)	(1.0005)	(0.0008)
S.Alt.1	147,367.9	0.2062	146,576.8	0.2057	1.0053	0.0004
S.Alt.2	116,955.8	0.1596	109,910.5	0.1490	1.0641	0.0105
S.Alt.3	119,256.4	0.1587	117,503.1	0.1569	1.0149	0.0018
S.Alt.4	195,999.4	0.2726	195,315.8	0.2723	1.0034	0.0002
R.Main	261.65	0.1503	252.39	0.1441	1.0368	0.0061
	(161.00)	(0.1102)	(154.50)	(0.1083)	(1.0420)	(0.0019)
R.Alt.1	305.64	0.1757	298.09	0.1710	1.0253	0.0047
R.Alt.2	272.68	0.1570	265.95	0.1528	1.0253	0.0042
R.Alt.3	262.38	0.1513	255.41	0.1457	1.0272	0.0055
R.Alt.4	380.33	0.2232	376.34	0.2198	1.0106	0.0033

Notes: Out-of-sample (OoS) and Within-sample (WS) mean absolute errors (MAEs) and mean relative errors (MREs) are computed for units for rent and sale following formula (2). The smaller MRE and MAE, the better the predictive power. We relate the two measures to each other by computing ratios MAE(OoS)/MAE(WS) for absolute measures and distances MRE(OoS)-MRE(WS) for relative measures reported in the last two columns. We also include in brackets figures on the median relative errors for the main models to give a clearer picture of the shape of the error curves.
Source: Authors' calculations based on data from the *Housing Observatory*, on advertisements available between 1 January and 31 December 2018.

3.2 The Magnitude, Direction and Source of Deviations

Every owner-occupier interviewed in the survey provides a hypothetical sales *and* a hypothetical rent price. We assess the differences between subjectively reported and imputed current values for owner-occupier. We also compare the deviations observed for sales prices and rent prices.

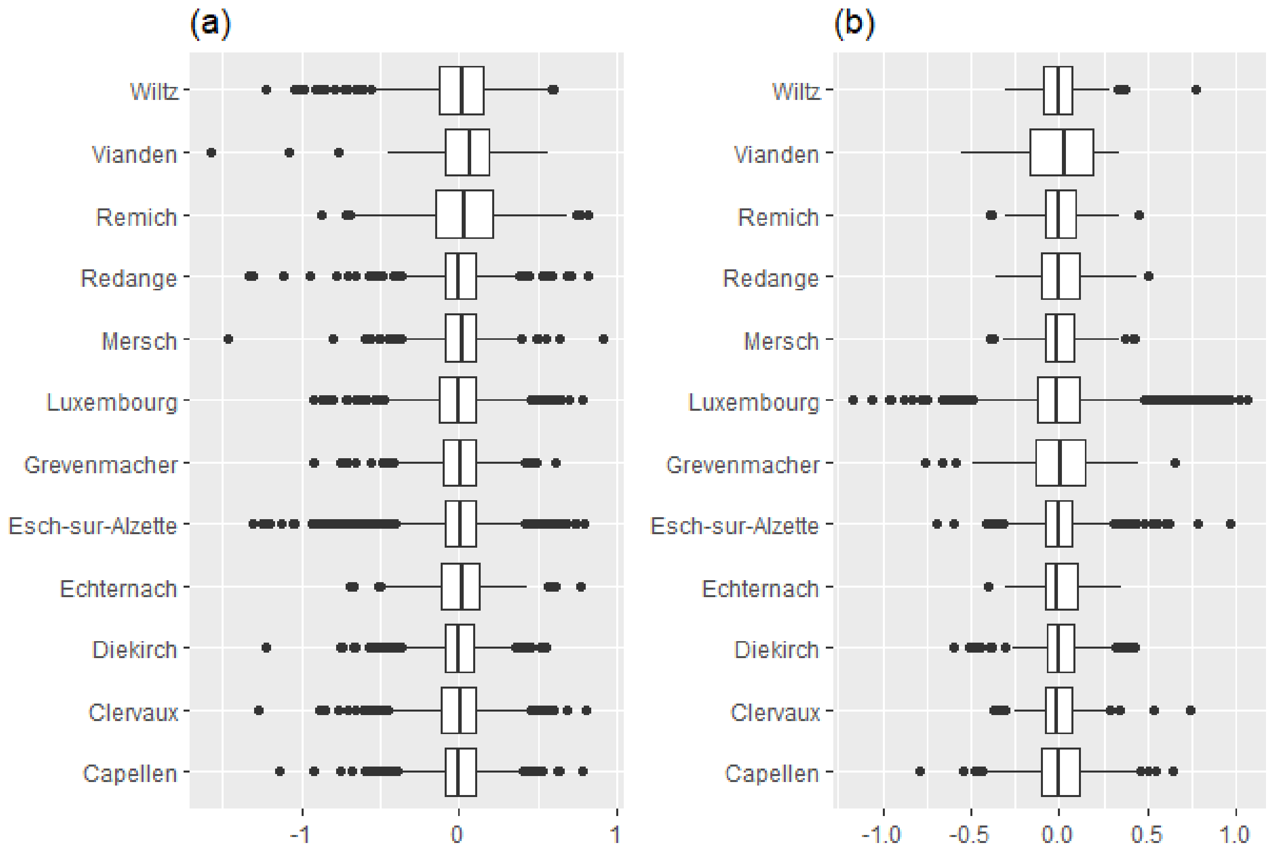
To obtain a first glimpse, we look at this relationship without any further controls: We find a positive Pearson correlation coefficient of $\rho = 38.13\%$ between the two types of deviations and Figure 4 visualises this strong link as a scatter plot. This means a smaller deviation of reported sales prices goes hand in hand with a smaller deviation of reported rent prices. Thus, survey respondents tend to be simultaneously good (or bad) in estimating hypothetical sales and rent prices revealing a certain measurable "degree of ability" performing such tasks in general.

Next, we characterise survey participants driving deviations making use of the large amount of additional information the survey provides. We therefore estimate multinomial logistic regressions describing the type of deviation via several sets of thematically grouped explanatory variables.

As response variable we classify survey participants by the direction of deviation of their reported prices from the imputed one. We label a value as *over-reported (under-reported)* if it is at least $x\%$ higher (lower) than the imputed counterpart. We consider deviations by less than $x\%$ as accurate and use this outcome as the reference category. As thresholds, we select $x \in \{10, 15\}$.

The error margins are selected to be in line with the relative out-of-sample errors measured via formula (2) and reported in the second column of Table 4: As we need to impute market prices and rents for dwellings in the survey, we need to choose a wide enough window to cover imputation uncertainty.

Figure 3 – Residuals by Canton.



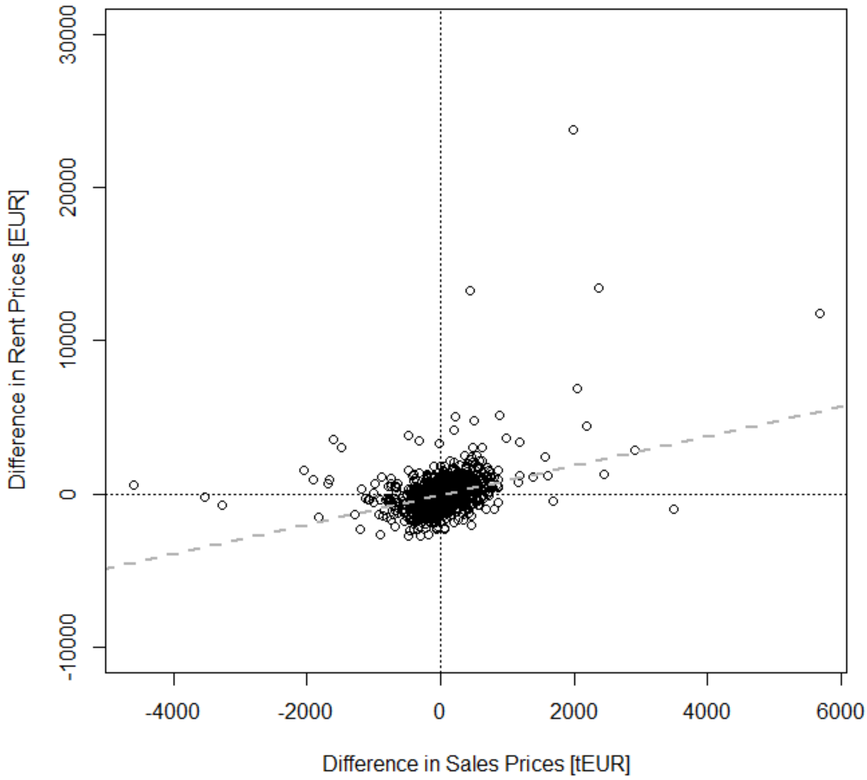
Notes: The figure depicts residuals clustered by canton retrieved from the S.Main (a) and R.Main (b) models.

For this, the out-of-sample error is relevant as dwellings appearing in the survey are usually not on the market as they are main residences at the moment of the interview. The total median out-of-sample relative error amounts to roughly 10% for both $S.Main$ and $R.Main$. The median OoS relative errors separately computed for positive and negative deviations are very similar. Precisely, we find median positive OoS errors for sales (rent) prices of 0.111 (0.114) and negative ones of 0.102 (0.105). We thus can assume a symmetric interval but need to select at least $x \geq 10\%$. To limit the probability that the deviation we measure and interpret here as reporting errors may in fact solely reflect hedonic estimation uncertainty, we also repeat the analysis for the even more conservative choice of $x = 15\%$.

To limit the amount of imposed assumptions, we do not use exact EUR-amounts of deviations as response variables, but rely on broad classifications differing between "accurate" estimates as well as positive and negative deviations. In a perfectly impartial allocation to deviation type, we would expect that the likelihood being classified as an accurate estimation was somewhat proportional to the length of the classification window, i.e., to $x \pm 10\%$ and $x \pm 15\%$. Furthermore, we assume that over- and under-reporting is *per se* equally likely, which we confirmed empirically before. These two technical assumptions imply a discrete uniform distribution on three classes with associated probabilities

$$\begin{aligned} \mathbb{P}(\text{Over-Reporting}) &= \frac{1-2x}{2}, \\ \mathbb{P}(\text{Accurate}) &= \quad 2x \qquad \text{and} \\ \mathbb{P}(\text{Under-Reporting}) &= \frac{1-2x}{2}. \end{aligned}$$

Figure 4 – Rent vs. Sales Price Deviation.



Notes: The figure compares the deviation between reported and imputed sales prices to the deviation between reported and imputed rent prices for each owner-occupied dwelling in the survey. The dashed line corresponds to a fitted linear regression model.

Hence, for a threshold of 10% (15%), we should expect 20% (30%) of all reported prices being accurate and 40% (35%) each being over- and under-reported.

Table 5 reports the theoretical and observed frequencies of deviation types. Generally, under-reporting is very common: We observe an excess number of under-reporters of roughly 10pp. In contrast, the number of over-reporters is much lower than expected.

Table 5 – Frequency of Deviation Types

	10% threshold		15% threshold	
Over-reporting	26.7 of 40%	(-13.35pp)	22.4 of 35%	(-16.48pp)
Accurate	24.4 of 20%	(+4.37pp)	35.4 of 30%	(+5.42pp)
Under-reporting	49.0 of 40%	(+8.99pp)	42.2 of 35%	(+7.23pp)

Notes: Over-reporting (under-reporting) means that the reported value exceeds (undercuts) the imputed one by $\pm x\%$. We call a reported price accurate whenever it does not deviate by more than $x\%$ from the imputed one. The frequency is expressed as percentage shares (%) in comparison to the *theoretically* expected share when assuming evenly distributed deviations. In parenthesis, we report the difference between observed and theoretical shares in percentage points (pp).

We assess the odds of the deviation exceeding $\pm x\%$ with respect to two groups of predictors: characteristics of the reference person of the household (H) and characteristics of the dwelling (HMR) both physical and locational.

We formalise this using the following logistic specification

$$\log\left(\frac{\mathbb{P}\left(\text{Over-reporting}|H, HMR\right)}{\mathbb{P}\left(\text{Accurate}|H, HMR\right)}\right) = \alpha^O + \beta_H^O H + \beta_{HMR}^O HMR, \quad \text{and} \tag{3}$$

$$\log\left(\frac{\mathbb{P}\left(\text{Under-reporting}|H, HMR\right)}{\mathbb{P}\left(\text{Accurate}|H, HMR\right)}\right) = \alpha^U + \beta_H^U H + \beta_{HMR}^U HMR. \tag{4}$$

Table 6 reports main estimation results with additional details listed in Table O.8 in the appendix. Appendix O.3 provides detailed descriptions of non-standard predictors entering the model (i.e., specific national questions not included in the harmonised European HFCS questionnaire).

Regarding HMR characteristics, we find greater deviations (both for under- and over-reporting) for homes acquired by the household long ago, suggesting that it becomes more difficult for owners to track and mentally adjust overall changes in house prices to a specific dwelling over extended periods of time. Some long-term owners not intending to sell their home may also simply consider this mentally challenging task as not being worth the effort.

Furthermore, under-reporting is more likely for larger premises. This holds true when measuring living space in square meters or using the type of HMR (houses versus apartments).

Some assessments by survey interviewers add valuable information:[13] Under-reporting current prices tends to be less likely when the interviewer reports a positive impression of interior or exterior conditions (i.e., dwellings rated as luxury or upscale from the exterior). This suggests that these ratings provide useful information.

Over-reporting is less common in the urban agglomerations Luxembourg City and Esch-sur-Alzette than in smaller municipalities and rural areas. However, under-reporting is more common in Luxembourg City than in the rest of the country, while it is less common in Esch-sur-Alzette.

The tenure status, renting versus owner-occupying, turns out to be a prime explanatory predictor: Renters tend to be much more likely to both under-report and over-report the market value of their home and less likely to report prices close to the ones suggested by the hedonic model. Households at the top of the net wealth distribution are more likely to over-report and less likely to under-report relative to imputed values. Likewise, higher household income is associated with a lower likelihood to mis-estimate.

In terms of education, we document a slightly lower likelihood to mis-report for persons with medium or high formal educational attainment, as compared to a low education level. However, the estimated coefficients are only significant for under-reporting and a medium education level.

4 Implications for Macroeconomic Statistics

4.1 Net Wealth, Residential Housing Wealth and its Distribution

Switching from reported survey data to imputed market values means an adjustment of the value of the current residential housing stock as well as measured net wealth at *current market prices*.

These computations may be particularly helpful for filling gaps in Luxembourg's canon of official statistics. STATEC, Luxembourg's national statistical institute, conducts an annual survey to update statistics on the current housing stock. A questionnaire is sent to every registered owner of newly constructed dwellings. However, only construction costs are collected,

Table 6 – The Source of deviations between reported and imputed values.

Over-reporting

	(1)		(2)		(3)		(4)		(5)	
Response: reported > imputed by	10%	15%	10%	15%	10%	15%	10%	15%	10%	15%
Intercept	1.091	0.642***	0.721	0.376*	0.489**	0.238***	1.354***	0.806**	1.032	0.521
	(0.092)	(0.051)	(0.398)	(0.197)	(0.177)	(0.087)	(0.143)	(0.084)	(0.642)	(0.323)
Housing status										
Owner					*(ref. cat.)*					
Renter	1.017	0.893	1.713*	1.533	2.577***	3.043***	1.064	0.948	2.571**	2.975**
	(0.212)	(0.192)	(0.542)	(0.471)	(0.883)	(1.064)	(0.227)	(0.208)	(1.091)	(1.291)
Education level										
Low (ISCED $\in \{0,1,2\}$)					*(ref. cat.)*					
Medium (ISCED $\in \{3,4\}$)			0.903	0.855					0.931	0.884
			(0.219)	(0.194)					(0.231)	(0.202)
High (ISCED $\in \{5-8\}$)			0.908	0.851					1.031	1.002
			(0.248)	(0.203)					(0.288)	(0.246)
Net income										
Q1					*(ref. cat.)*					
Q2			0.650	0.732					0.630	0.722
			(0.240)	(0.332)					(0.240)	(0.343)
Q3			0.443*	0.389*					0.436	0.382
			(0.196)	(0.196)					(0.206)	(0.203)
Q4			0.440**	0.484**					0.424**	0.455**
			(0.146)	(0.168)					(0.150)	(0.167)
Q5			0.528	0.534					0.518	0.516
			(0.202)	(0.248)					(0.210)	(0.250)
Net wealth										
Q1					*(ref. cat.)*					
Q2			1.608	1.286					1.769	1.448
			(0.746)	(0.532)					(0.845)	(0.627)
Q3			1.155	0.929					1.290	1.031
			(0.538)	(0.400)					(0.611)	(0.459)
Q4			1.852	1.565					2.040	1.644
			(0.894)	(0.644)					(1.012)	(0.705)
Q5			3.753***	3.532***					4.570***	4.133***
			(1.770)	(1.561)					(2.287)	(1.942)
Type of HMR										
House					*(ref. cat.)*					
Apartment					0.789	0.617**			0.774	0.615**
					(0.190)	(0.142)			(0.201)	(0.149)
Surface of HMR					1.001	1.000			0.998	0.998*
					(0.001)	(0.001)			(0.002)	(0.001)
Years since acquisition [log]					1.337***	1.418***			1.170	1.261*
					(0.134)	(0.148)			(0.146)	(0.162)
Interviewer Rating: exterior conditions										
mid-range, modest or low-income					*(ref. cat.)*					
luxury or upscale					1.254	1.507***			1.247	1.482**
					(0.195)	(0.239)			(0.202)	(0.253)
Interviewer Rating: interior conditions										
good, fair, poor or										
dwelling not seen by the interviewer					*(ref. cat.)*					
excellent					0.847	0.882			0.824	0.869
					(0.188)	(0.200)			(0.195)	(0.211)
Canton										
Countryside					*(ref. cat.)*					
Luxembourg City							0.682*	0.599**	0.524***	0.458***
							(0.150)	(0.124)	(0.124)	(0.109)
Esch sur Alzette							0.614***	0.647***	0.640**	0.668**
							(0.107)	(0.107)	(0.121)	(0.118)

Table 6 – The Source of deviations between reported and imputed values (cont'd).

Under-reporting

	(1)		(2)		(3)		(4)		(5)	
Response: reported < imputed by	10%	15%	10%	15%	10%	15%	10%	15%	10%	15%
Intercept	1.505***	0.865*	4.349***	2.843**	0.980	0.447**	1.480***	0.873	1.357	0.656
	(0.121)	(0.068)	(1.943)	(1.165)	(0.310)	(0.147)	(0.162)	(0.086)	(0.734)	(0.345)
Housing status										
Owner					*(ref. cat.)*					
Renter	3.273***	3.385***	1.719**	1.670**	3.275***	4.434***	3.155***	3.301***	2.422***	3.649***
	(0.562)	(0.518)	(0.434)	(0.398)	(0.878)	(1.203)	(0.545)	(0.513)	(0.817)	(1.213)
Education level										
Low (ISCED $\in \{0,1,2\}$)					*(ref. cat.)*					
Medium (ISCED $\in \{3,4\}$)			0.796	0.714*					0.783	0.707*
			(0.168)	(0.138)					(0.170)	(0.144)
High (ISCED $\in \{5-8\}$)			1.000	0.868					0.849	0.756
			(0.259)	(0.187)					(0.237)	(0.174)
Net income										
Q1					*(ref. cat.)*					
Q2			0.714	0.767					0.822	0.914
			(0.229)	(0.265)					(0.277)	(0.331)
Q3			0.465**	0.413**					0.562	0.500*
			(0.155)	(0.137)					(0.196)	(0.170)
Q4			0.321***	0.330***					0.383***	0.391***
			(0.096)	(0.098)					(0.124)	(0.128)
Q5			0.333***	0.311***					0.367**	0.344**
			(0.120)	(0.115)					(0.140)	(0.134)
Net wealth										
Q1					*(ref. cat.)*					
Q2			1.733*	1.569					1.634	1.473
			(0.570)	(0.481)					(0.549)	(0.494)
Q3			0.993	0.914					0.916	0.799
			(0.344)	(0.269)					(0.328)	(0.257)
Q4			0.576	0.510**					0.426**	0.333***
			(0.193)	(0.163)					(0.155)	(0.125)
Q5			0.521*	0.501*					0.253***	0.208***
			(0.201)	(0.181)					(0.112)	(0.090)
Type of HMR										
House					*(ref. cat.)*					
Apartment					2.781***	2.558***			2.519***	2.362***
					(0.583)	(0.421)			(0.611)	(0.430)
Surface of HMR					1.003**	1.002***			1.006***	1.006***
					(0.001)	(0.001)			(0.001)	(0.001)
Years since acquisition [log]					1.234**	1.342***			1.400***	1.686***
					(0.107)	(0.119)			(0.167)	(0.210)
Interviewer Rating: dwelling (exterior)										
mid-range, modest or low-income					*(ref. cat.)*					
luxury or upscale					0.539***	0.588***			0.641***	0.726*
					(0.082)	(0.091)			(0.107)	(0.120)
Interviewer Rating: interior conditions										
good, fair, poor or										
dwelling not seen by the interviewer					*(ref. cat.)*					
excellent					0.475***	0.481***			0.540***	0.563***
					(0.096)	(0.075)			(0.116)	(0.098)
Canton										
Countryside					*(ref. cat.)*					
Luxembourg City							1.399**	1.262	1.970***	1.802***
							(0.234)	(0.185)	(0.388)	(0.322)
Esch sur Alzette							0.780	0.767*	0.665**	0.620***
							(0.120)	(0.110)	(0.112)	(0.100)
No. of Observations	1,616	1,616	1,616	1,616	1,616	1,616	1,616	1,616	1,616	1,616
McFadden's Pseudo-R^2	0.029	0.034	0.101	0.112	0.074	0.082	0.039	0.042	0.154	0.169
AIC	3,289	3,338	3,116	3,138	3,161	3,195	3,265	3,317	2,967	2,974
BIC	3,311	3,359	3,320	3,343	3,248	3,281	3,308	3,360	3,257	3,265

Notes: We model deviations between reported and imputed current market vales of the household main residence. Additional controls are reported in Table O.8.
Models (1) – (5) refer to distinct specifications. We distinguish two response variables describing the deviation between imputed and reported current values of the HMR. Over-reporting and under-reporting, respectively, means that the reported value diverges by ±10% or ±15% from the imputed one.
$Q1$ to $Q5$ denote first to fifth quintiles. Goodness-of-fit measures report the average over five implicates. Significance is indicated using standard notation: *p-value<0.1; **p-value<0.05; ***p-value<0.01. The reported Pseudo-R^2 is computed following McFadden (1974).
Source: LU-HFCS, 3rd wave and authors' calculations.

which are usually far from market sales prices.[14] In addition, in the non-financial accounts dwellings (ESA, 2010, code: AN.111) and land underlying buildings and structures (ESA, 2010, code: AN.2111) are currently not distinguishable between households and non-profit institutions serving households (NPISHs). There are no reliable estimates of the current market value of the residential housing stock owned by private households (see also EG-LMM, 2020).

We define net wealth for household h and time t as

$$\text{Net Wealth}_{ht} = \text{Real Assets}_{ht} + \text{Financial Assets}_{ht} - \text{Liabilities}_{ht}. \tag{5}$$

Real assets include housing assets, financial assets exclude public and occupational pensions, and liabilities comprise mortgages and non-mortgage debt (see Chen et al., 2020, for details).

Table 7 reports the residential housing wealth and net wealth of owner-occupiers according to reported and imputed market values for different quintiles of the income and wealth distribution. We focus here on *distributional indicators* revealing the impact on different income and wealth groups using objectively imputed values rather than subjectively reported values.

Table 7 – Distributional Wealth Measures.

	Shares			Median [EUR]			
	Reported [%]	Imputed [%]	Difference [pp]	Reported	Imputed	Difference	
Complete Sample ($N = 1,616$, $N_w = 226,378$)							
Residential Housing Wealth							
Total	100	100	–	600,000	654,104	54,104	***
				(14,277)	(8,129)		
Net Wealth							
Total	100	100	–	498,454	563,334	64,880	
				(23,399)	(20,047)		
Net Wealth Breakdowns by							
Net Income – Q1	8.8	9.6	0.9	74,210	79,292	5,082	
Net Income – Q2	12.2	12.7	0.5	272,193	389,896	117,703	
Net Income – Q3	14.0	14.4	0.5	473,780	549,618	75,838	
Net Income – Q4	20.9	20.9	0.0	707,394	746,932	39,538	
Net Income – Q5	44.2	42.3	-1.9	1,040,763	1,013,534	-27,229	
Net Wealth – Q1	0.2	0.2	0.0	7,060	7,736	676	
Net Wealth – Q2	3.9	5.5	1.6	157,138	203,639	46,501	
Net Wealth – Q3	11.1	12.6	1.5	498,751	558,022	59,271	***
Net Wealth – Q4	19.0	19.4	0.4	839,520	852,966	13,446	
Net Wealth – Q5	65.8	62.3	-3.5	1,858,008	1,840,418	-17,590	**
Owner-Occupiers ($N = 1,207$, $N_w = 156,210$)							
Residential Housing Wealth							
Total	100	100	–	652,000	696,994	44,994	***
				(18,730)	(11,731)		
Net Wealth							
Total	100	100	–	732,360	783,653	51,293	***
				(19,718)	(17,050)		
Net Wealth Breakdowns by							
Net Income – Q1	8.9	9.9	0.9	513,081	683,383	170,302	***
Net Income – Q2	11.7	12.2	0.5	573,937	646,734	72,797	***
Net Income – Q3	14.0	14.5	0.5	636,916	702,271	65,355	***
Net Income – Q4	20.8	20.9	0.0	784,495	804,944	20,448	**
Net Income – Q5	44.6	42.6	-2.0	1,121,400	1,111,741	-9,659	
Net Wealth – Q1	0.0	0.0	0.0	2,240	56,989	54,749	*
Net Wealth – Q2	2.9	4.6	1.7	226,010	326,988	100,977	***
Net Wealth – Q3	11.1	12.6	1.5	508,324	570,284	61,959	***
Net Wealth – Q4	19.7	20.1	0.4	844,560	858,481	13,921	*
Net Wealth – Q5	66.4	62.7	-3.7	1,849,368	1,823,905	-25,463	***

Notes: Residential Housing Wealth refers to the total current value of households' main residences (neglecting partial ownership when owners may not all belong to the same household). N refers to the number of individual survey observations and N_w the sum of survey weights, thus, the number of households they represent in Luxembourg. Net Wealth follows the definition in equation (5) taking partial ownership of residential housing wealth into account. $Q1$ to $Q5$ denote quintiles. Standard deviations are reported in parentheses. Asterisks indicate significant differences between imputed and reported values. Test statistics are based on median quantile regressions taking into account replicate weights and the multiply imputed nature of the LU-HFCS. Significance is coded using standard notation: *p-value<0.1; **p-value<0.05; ***p-value<0.01. Totals in EUR are reported in Table O.4.
Source: LU-HFCS, 3rd wave and authors' calculations based on data from the *Housing Observatory*, on advertisements available between 1 January and 31 December 2018.

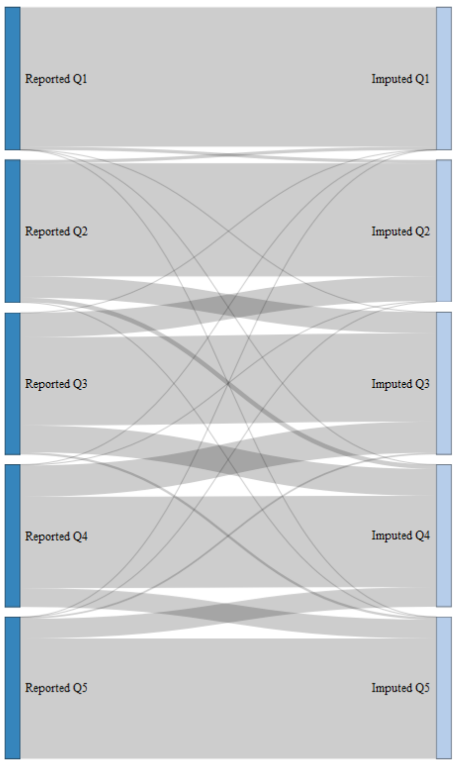
Overall, our imputed values increase average total residential housing wealth – and thus also total net wealth – held by Luxembourg households. The increase is substantial in magnitude and also significant in statistical terms: imputed market values increase total median net worth of owner-occupiers by roughly EUR 50,000.

We observe the largest increases among households belonging to the second net wealth quintile. For the median owner-occupiers in this quintile, net wealth increases by around EUR 100,000. Using reported values, the lowest quintile of households held approximately zero net wealth at the median, i.e., the current value of their liabilities roughly match the current value of their assets. With imputed prices, this number becomes positive.

According to the imputed values, the top quintile (both in terms of income and wealth) are less prosperous than they report in surveys. This result may stem from our imputation model that – though controlling for a wide range of housing characteristics, including exceptional equipment – has a regression-to-the-mean tendency, which may miss certain characteristics neither captured by the survey nor the market data.

Figure 5 analyses the changes in the net wealth distribution resulting when switching from reported to imputed prices: where the change has no effect, observations stay in the same quintile. In other words, there would only be zeros off the diagonal in the transition matrix in Table 8. While low-wealth households indeed largely remain in the lowest wealth quintile, a non-negligible share of households are kicked out of the second (4 on 20), third (8 on 20), fourth (7 on 20) and top (3 on 20) quintile. Although switches predominantly occur between neighbouring quintiles, a small fraction also moves to more distant groups.

Figure 5 – Changes in Net Wealth Distributions



Notes: For the complete sample (owners and renters), the figure shows the number of observations where the imputations lead individual households to shift across quintiles in the net wealth distribution.

Overall, these changes in total net wealth lead to a small decrease of measured wealth inequality along the income and wealth distribution as the distributions become more narrow.

Table 8 – Changes in Net Wealth Distribution

		Imputed					
		Q_1	Q_2	Q_3	Q_4	Q_5	Total
Reported	Q_1	19.65	0.42	0.00	0.00	0.00	20.07
	Q_2	0.41	16.00	2.97	0.58	0.08	20.04
	Q_3	0.00	3.41	12.44	3.86	0.23	19.94
	Q_4	0.00	0.11	4.32	12.84	2.75	20.02
	Q_5	0.00	0.02	0.29	2.72	16.91	19.94
Total		20.06	19.97	20.01	20.00	19.96	100.00

Notes: For the complete sample (owners and renters), the table reports the percent of observations changing quintile in the net wealth distribution due to the imputation. Figure 5 presents these results graphically.

These quite significant changes along several dimensions indicate that an objectified evaluation does not simply shift the entire distribution but rather reveals differential effects along several dimensions. Such a finding can lead to biases in wealth decompositions such as, for instance, distributional national accounts (see Waltl and Chakraborty, 2022; EG-LMM, 2020; Waltl, 2022).

Such changes are likely to be larger for countries with a significant share of owner-occupiers. For purely illustrative purposes, applying the shifts along the wealth and income distribution identified in Luxembourg to other countries participating in the HFCS provides an indication of how important reporting issues could be in this domain.

For this hypothetical analysis of other countries, we take country-specific total residential housing wealth and owner-occupation rate, and apply Luxembourg's measured mis-reporting rates per wealth and income[15] quintile to owner-occupiers in other countries. To obtain a full wealth measure, we then subtract reported residential housing wealth from total wealth and plug in the adjusted equivalent. To obtain these adjusted wealth distributions, the following steps are applied.

Denoting reported residential housing wealth in net wealth (gross income) quintile WQ_i (IQ_i), $i \in \{1, \ldots, 5\}$, by HW_{WQ_i} (HW_{IQ_i}) and its imputed counterpart by $HW^I_{WQ_i}$ ($HW^I_{IQ_i}$), hypothetical figures for other countries are computed using the adjustment factors

$$\alpha_{WQ_i} = \frac{HW^I_{WQ_i} - HW_{WQ_i}}{HW_{WQ_i}} \quad \text{and} \quad \alpha_{IQ_i} = \frac{HW^I_{IQ_i} - HW_{IQ_i}}{HW_{IQ_i}}. \tag{6}$$

For each wealth and income quintile, the adjusted totals are given by

$$RHW^{adj} = \sum_{i=1}^{5} \alpha_{WQ_i} HW_{WQ_i} = \sum_{i=1}^{5} \alpha_{IQ_i} HW_{IQ_i}. \tag{7}$$

Total quintile-specific net wealth is then obtained by replacing reported housing wealth and by the adjusted figures as defined in (7),

$$\text{Net Wealth}^{adj}_{WQ_i} = \text{Net Wealth}_{WQ_i} - RHW^{obs}_{WQ_i} + RHW^{adj}_{WQ_i} \quad \text{and} \tag{8}$$

$$\text{Net Wealth}^{adj}_{IQ_i} = \text{Net Wealth}_{IQ_i} - RHW^{obs}_{IQ_i} + RHW^{adj}_{IQ_i}. \tag{9}$$

Table 9 reports the adjustment factors α_{IQ_i} and α_{WQ_i} found for Luxembourg and defined in (6). Applying these to the country-specific totals for residential housing wealth (following the same definitions as in Luxembourg) yields hypothetical distributional impacts for each country.

Table 9 – Distributional Wealth Measures: Owner-occupiers.

	Shares			Totals	
	Reported [%]	Imputed [%]	Change α [pp]	Reported [mEUR]	Imputed [mEUR]
Residential Housing Wealth					
Total	100.000	105.694	5.694	116,713	123,358
Breakdowns by					
Gross Income – Q1	10.165	12.590	2.425	11,864	14,694
Gross Income – Q2	12.502	14.108	1.606	14,592	16,466
Gross Income – Q3	18.456	19.967	1.511	21,540	23,304
Gross Income – Q4	24.402	25.155	0.753	28,481	29,360
Gross Income – Q5	34.473	33.873	-0.600	40,235	39,535
Net Wealth – Q1	0.795	0.898	0.103	928	1,048
Net Wealth – Q2	8.921	12.093	3.172	10,412	14,114
Net Wealth – Q3	19.459	22.741	3.281	22,712	26,541
Net Wealth – Q4	28.134	29.854	1.720	32,836	34,843
Net Wealth – Q5	42.689	40.108	-2.582	49,824	46,811

Notes: In the Luxembourg HFCS (3rd wave) there are $N = 1,207$ survey observations representing $N_w = 156,210$ households. Changes in percentage points are computed following formula (6).
Source: LU-HFCS, 3rd wave and authors' calculations based on data from the *Housing Observatory*, on advertisements available between 1 January and 31 December 2018.

This procedure implicitly assumes that owner-occupiers in other countries are prone to reporting issues of a similar relative magnitude to those in Luxembourg. More precisely, the deviation between reported and imputed values is assumed to be proportional to either household gross income or net wealth and not *systematically* related to any other observable factor.

Results are compiled by first adjusting quintile-specific wealth for owner-occupiers. The proportion of owner-occupiers varies widely across countries in the European HFCS (ranging from 43.9% in Germany to 88.8% in Slovakia in 2018). Indeed, the relation between owner-occupation rate and the change in net wealth is statistically significant as reported in Table 10.

Therefore, we select four countries participating in the HFCS and representing different realities in this regard (see Table 10 for details). This includes the countries with the lowest (Germany) and the highest (Slovakia) proportion of owner-occupiers, as well as two typical intermediate countries (France and Italy). We expect changes to be larger for countries with a larger share of owner-occupiers.

Table 10 – As-if analysis

	Germany	France	*Luxembourg*	Italy	Slovakia
Owner-occupation Rate [%]	43.9	57.9	*69.0*	68.5	88.8
Change in Net Wealth [%]	3.28	3.49	*3.56*	3.89	4.20
ρ	93.76% (p-value: 0.0185)				

Notes: The table reports changes in measured total net wealth for four European countries increasingly ordered by the change in net wealth when applying mis-reporting shares found for Luxembourg. ρ measures the Pearson correlation between changes in measured total net wealth and the owner-occupation rate across countries. The p-value reports results of a *Pearson's product-moment correlation test* for $H_0: \rho = 0$ rejected at the 5% level.
Source: HFCS, 3rd wave.

Table O.5 in the appendix reports the full set of results. In general, the adjustment factors computed for Luxembourg lead to a *small decrease in inequality* and a *large increase in overall* *wealth*. As reported in Table 10, total net wealth is most affected in Slovakia, which also had the largest share of owner-occupiers. At the other extreme, total net wealth is least affected in Germany, which had the lowest share of owner-occupiers. This finding suggests that wealth distributions may be poorly measured in countries with large shares of owner-occupiers or shares of residential housing in overall. Given the very large amounts at play even small shifts – whenever found to be systematic – imply large aggregate impacts.

4.2 Housing Market Indicators

Common housing market measures include the *price-to-rent* (PR) *ratio*, *price-to-income* (PI) *ratio*, and the *rent-to-income* (RI) *ratio*. PI and RI ratios are widely used affordability measures. A high RI ratio implies households spend a large share of monthly earnings on rent payments. A high PI ratio suggests it is more difficult to purchase a home.

The PR ratio is considered a measure of investment potential and the sustainability of the housing market: a high ratio suggests that sales prices are high compared to potential returns, i.e., rents. This implies high expected future house prices according to the user cost model for durable goods (Hicks et al., 1975). For residential housing, Himmelberg et al. (2005) show that in equilibrium the cost of buying and using a housing unit for a period (the sales price P_t multiplied by the per-dollar user cost u_t) should equal the total rent for the same period (R_t):

$$R_t = u_t \cdot P_t, \tag{10}$$

where u_t includes the interest rate, maintenance and average transaction costs, the depreciation rate for housing, a risk premium for owning as opposed to renting, and expected capital gains over the period. In the case of Luxembourg, mortgage payments are tax deductable (see subsection 4.3), which is also accounted for in the user cost formula. This framework implies that in periods of high PR ratios, investing in properties is less attractive due to a limited earning potential.

Hill and Syed (2016) observe that the PR ratio is frequently monitored over time because a sales price index and a rent price index are the only inputs needed for its calculation. However, direct measurement of the PR ratio is rare because it is hard to simultaneously observe the price and the rent for a given dwelling. However, our double-imputed data set allows us to calculate both sales price and rent for each dwelling and match them to the income of the current inhabitant.

There are two main challenges for measuring PR ratios: Commonly used measures either fail to be (*i*) representative for the entire stock of houses or (*ii*) compare prices of very heterogeneous properties (see Bracke, 2015; Hill and Syed, 2016; Waltl, 2018, for procedures to compute quality-adjusted ratios when relying directly on market data).

Issue (*i*) arises when relying on transaction data only, as this usually represents just a tiny share of the housing stock and is unlikely to be representative for the housing stock as a whole. Issue (*ii*) is particularly important to obtain unbiased ratios as the equilibrium condition in the user cost model implicitly assumes that P_t and R_t refer to properties of equivalent quality as noted by Hill and Syed (2016).

A third issue (*iii*) arises when income enters as additional dimension to compile RI or PI. For market sales or rent price data matching information on household income is hardly ever available. It is thus cumbersome to compare income to prices or rents beyond comparing the median (see also Gan and Hill, 2009). In our case, the household inhabiting each dwelling has reported its total household income.

The double-imputed data overcomes all three types of shortcomings and allows us to compute

Figure 6 – Prices versus Rents versus Income

Notes: The figures depict PR, PI and RI ratios across their respective distributions relying on imputed prices and rents. The bars indicate a 90% confidence interval. Ratios are depicted separately for renters, owner-occupiers and the combined sample. The corresponding values are reported in Table O.9.

un-biased population-representative ratios by aggregating *individual ratios*. Thus, no kind of quality-adjustment a la Hill and Syed (2016) is needed to ensure a comparison of like with like and aggregation issues are circumvented.

Concretely, we relate imputed sales prices to imputed annual rents and annual household income. We therefore report population-representative indicators:

$$PR(\vartheta) = 100 \cdot Q_\vartheta \left(\frac{P_{ht}}{12 \cdot R_{ht}} \right), \tag{PR}$$

$$PI(\vartheta) = 100 \cdot Q_\vartheta \left(\frac{P_{ht}}{12 \cdot I_{ht}} \right), \tag{PI}$$

$$RI(\vartheta) = 100 \cdot Q_\vartheta \left(\frac{R_{ht}}{I_{ht}} \right), \tag{RI}$$

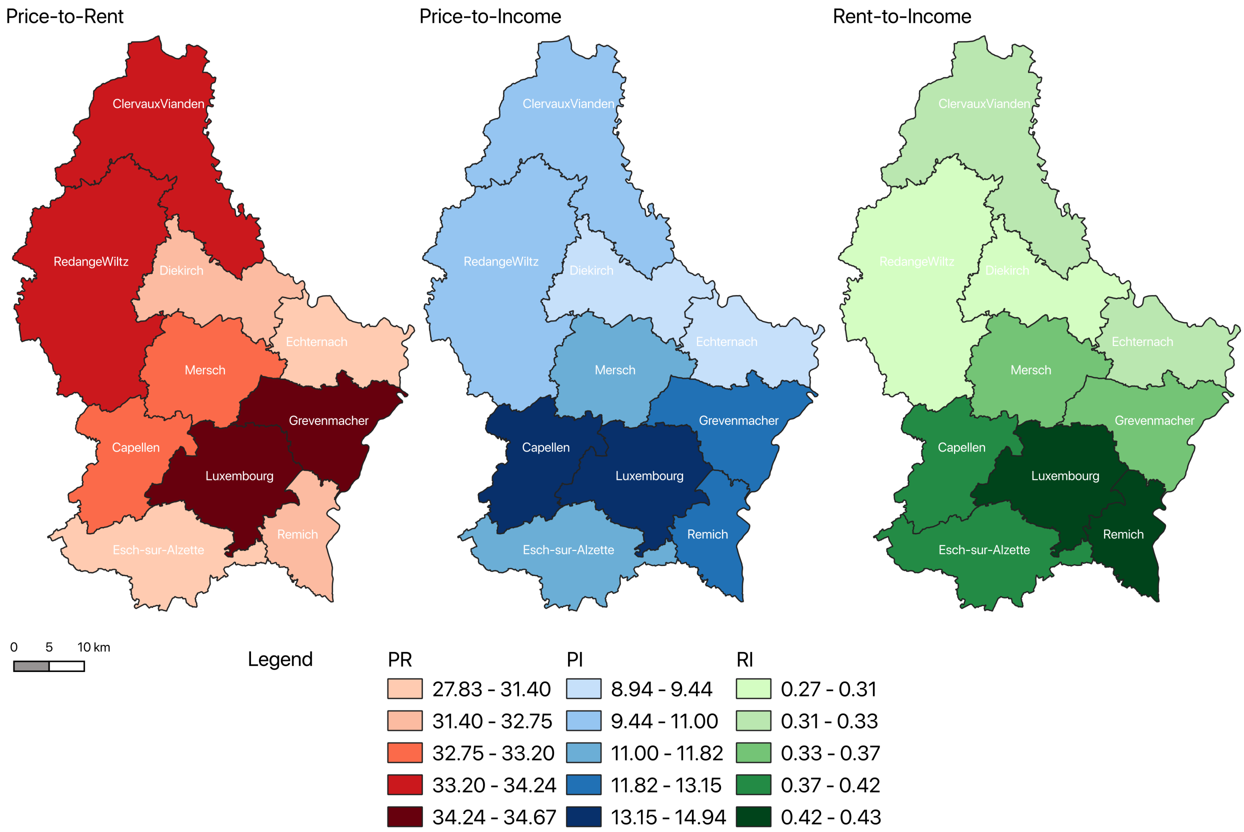
where $\vartheta \in \{10, 25, 50, 75, 90\}$ indexes the percentile of the distribution at which the the ratios are evaluated.

Results appear in Figure 6. At the median, PR ratios are about 32. In international[16] comparisons, such a level is considered high and associated with low sustainability. The results also reveal that owner-occupiers have higher PR ratios over large parts of the income distribution. These differences are often statistically significant, as indicated by the non-overlapping confidence intervals.

In line with our argumentation, the OECD writes that aggregate statistics "provide only a general indication of the extent to which housing is (un)affordable for a (median) household, they are ill suited to support policy makers in targeting housing supports to different groups" (OECD, 2021, Box 1.1). Our micro-data yields more disaggregated indicators revealing a substantial degree of variation within Luxembourg (see Figure 7): differences appear across regions and across tenure types. Particularly the Canton including the capital Luxembourg City displays very high ratios indicating pronounced affordability concerns.

Similar results are obtained when comparing house prices to household income. For 80% of the population the PI ratio varies between 6.4 and 31.3. The median of 12 indicates that for 50% of the population acquiring their home at current market prices would require financial resources

Figure 7 – Regional Variation in Ratios

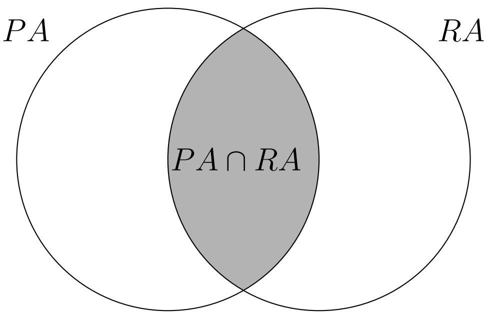


Notes: The figures depict median PR, PI and RI ratios by canton relying on imputed prices and rents (renters and homeowners). Cantons with a low number of survey respondents are merged with their neighbours. Corresponding values are reported in Table O.1 in Appendix O.1.

equal to 12 total annual household incomes – excluding transaction costs and interest. The PI ratio is higher for renters than for owner-occupiers, reflecting lower incomes among renters (see Table O.7 in Appendix O.4).

4.3 Housing Affordability Indicators

Linneman and Wachter (1989) show how borrowing constraints reduce the likelihood of buying a home. Low earnings reduce the repayment capacity (income effects) and insufficient savings limit a household's down-payment required to secure a mortgage (wealth effects). Gan and Hill (2009) operationalize income effects as *repayment affordability* (RA) and wealth effects as *purchase affordability* (PA). Both criteria have to be fulfilled for a sustainable home purchase as indicated by the grey-shaded area below.



PA is defined as a household's ability to finance the purchase. This criteria is fulfilled if either the household owns enough funds right away or is eligible to borrow sufficiently to purchase a home. RA, however, captures the ability to bear the monthly financial burden imposed on a household repaying its mortgage debt. A successful purchase takes place if both criteria are simultaneously fulfilled.

To assess affordability, we restrict the analysis to current renters, i.e., potential owners, and their financial situation. We assume the dwelling they are currently renting matches their current needs. This assumptions allows us to address the question: Could renters buy their main residence?

Table 11 – Calibration Parameters

Description		Rate / Value	Details
τ_h	Transaction Costs Components		in % of the purchase price
	Registration tax	6%	✓
	Transcript tax	1%	✓
	Surtax	3%	✓; Luxembourg City only
	Notary fees	≈1.5%	✗
τ^{C}	Total Transaction Costs	11.5%	within Luxembourg City
$\tau^{\neg C}$	Total Transaction Costs	8.5%	outside of Luxembourg City
η_h	Tax exempt	household-specific	granted tax exempt in % of the purchase price
i	Average lending rate	1.74%	can be deducted in the income tax declaration
m	Maximum maturity for mortgages	30 years	
$\max(a_h)$	Maximum age for repaying mortgages	66 years	

Notes: The table summarises key calibration variables for computing purchase and repayment affordability measures. The reported rates are as of 2018. The table specifies which taxes are eligible for the tax credit *Bëllegen Akt* (✓/✗).

Table 11 summarises Luxembourg-specific calibration parameters. Additionally, we adjust for policies affecting truly incurred costs as Luxembourg decreases the entry barriers for first-time owners via several favourable tax treatments (see LISER, 2022). These include most prominently the *Bëllegen Akt* providing first-time buyers (which we identify with renters in the HFCS) with a tax credit on certain purchase-related administrative fees and taxes up to a total of EUR 20,000 for a single buyer and EUR 40,000 for couples (see again Table 11 for details). Additionally, interest payments on mortgages are tax-deductible within the annual income tax declaration (see also Kaempff, 2018; Girshina et al., 2021). Further subsidies are granted depending on income and family situation. We do not take these into account, as one can only apply for them after the acquisition (thus not directly affecting PA) and they constitute one-time payments (thus not affecting repayment affordability).[17]

To assess *purchase affordability* (PA), we calculate the initial loan-to-value (LTV) ratio needed to finance the house purchase. P_h denotes the (imputed) current market value of household h's main residence and L_h the amount of mortgage taken out.

In December 2020, Luxembourg authorities issued a regulation requiring credit institutions to cap the LTV ratio at 100% for first-time buyers.[18] We thus consider a renting household h will fulfil the purchase affordability criteria if and only if

$$LTV(h) = \frac{L_h}{P_h} \leq 100\%. \tag{11}$$

Furthermore, we assume that households can use all their net liquid[19] assets for their down-payment. Net liquid assets NLA_h are defined as the sum of liquid financial assets FA_h (deposits, mutual funds, bonds, publicly traded shares and managed accounts) minus non-collateralised debt D_h (outstanding balance of credit lines/overdraft and credit card debt),

$$NLA(h) = FA_h - D_h. \tag{12}$$

A LTV ratio less than or equal to 100% implies that $NLA(h)$ must cover at least transaction costs τ_h. Around 0.7% of all renters have sufficient net liquid assets to finance the purchase without a mortgage. The required external financing is estimated to be EUR 562,000 for the average renter and EUR 508,000 for the median renter.

Transaction costs are rather high in Luxembourg.[20] Costs for the buyer include the registration tax (6% of the property price), the transcript tax (1%), notary fees (around 1.5%) and an additional surtax for Luxembourg City (3%). This means minimum transaction costs of 11.5% in Luxembourg City and 8.5% in the rest of the country. Real estate agents' commissions are not considered here as they are typically payable by the seller.

In Luxembourg, housing subsidies and tax benefits amount to 3.1% of government expenditures (Kaempff, 2018). First, mortgage interest payments qualify as deductions in the annual income tax declaration and thus affect RA (see Table 11 for details). Second, the most prominent policy is the *Bëllegen Akt* supporting first-time home buyers and affecting PA. It was introduced in 1990 and last reformed in 2002. In its current version, several fees and taxes are deductible to an absolute limit of EUR 20,000 (EUR 40,000 for jointly purchased homes). By assumption, all renters are first-time buyers and thus eligible for the full tax exemption. In total, this yields

$$NLA(h) \geq \tau_h(1 - \eta_h)P_h, \tag{13}$$

where $\eta_h \in [0; 1)$ is the household-specific tax exemption expressed as a share of the purchase price (see Table 11).

For 2018, we find that 69.4% (3.0 pp)[21] of all renting households do not *jointly* fulfil purchase affordability criteria (11) and (13) meaning they do not have sufficient net liquid assets for the required transaction costs. Only 5.7% (1.4 pp) of all renting households could achieve an initial LTV of 80% or less when buying their main residence at market prices.

Next, we assess RA by calculating the monthly amount of available income. Assessing households' entire financial situation allows us to identify whether they would need a mortgage to finance the purchase of their current home and include the implied additional spending into this assessment. For that purpose, we define a *financial margin* (FM) for each household h as its net income NI_h minus debt payments D_h minus basic living costs CoL_h. If a household has a non-negative margin, we classify it as fulfilling the repayment affordability criteria,

$$FM(h) = NI_h - D_h - CoL_h \geq 0. \tag{14}$$

Households report their net income directly in the LU-HFCS. Debt payments consist of two distinctive parts: (i) non-collateralised debt payments and payments for other mortgages (collected in the LU-HFCS); (ii) imputed payments for the mortgage required to purchase the dwelling at current market prices.

The annual payments D_h related to part (ii) are the product of the initial debt and an annuity factor α_h,

$$\alpha_h = \frac{i \cdot (1+i)^{m_h}}{(1+i)^{m_h} - 1}, \tag{15}$$

where $m_h \in \mathbb{N}$ denotes the maximum maturity in years applicable to household h and $i > 0$ the agreed interest rate. We assume an interest rate of $i = 1.74\%$, which was the average annualised agreed lending rate for house purchases in 2018.[22] We estimate the initial debt for renters as the market value of their main residence P_h plus transaction costs τ_h (taking into account any tax exemptions) minus net liquid assets $NLA(h)$.

As mentioned above, mortgage interest payments are tax deductible.[23] In the first five years, the deductible amount is limited to EUR 2,000 per year and household member (reference person, spouse or registered partner and their children).[24] The taxable income is not directly reported in the HFCS, so we need to estimate the amount saved via the tax-deductibility of interest payments relying on a strong assumption: we multiply the marginal tax rate of 39% (+ 7% of this amount as an additional contribution to the employment fund) by the amount of tax-deductible interest payments.[25] Furthermore, we subtract the amount of taxes saved via the tax deductibility from mortgage interest payments.

Table 12 – Affordability for Selected Household Categories

	PA ∩ RA		PA		RA	
	[%]	[pp]	[%]	[pp]	[%]	[pp]
Total	18.1	(2.2)	30.6	(3.0)	37.0	(2.7)
Age						
16-34	29.3	(5.3)	35.4	(5.5)	61.1	(5.9)
35-44	22.8	(5.0)	29.7	(6.8)	51.7	(6.0)
45-54	17.5	(4.6)	37.4	(6.8)	32.5	(5.8)
55-64	4.9	(2.4)	17.9	(5.1)	5.7	(2.7)
65+	1.4	(1.6)	31.8	(11.6)	1.4	(1.6)
Net Liquid Assets						
Q1	0.0	–	0.0	–	23.0	(4.3)
Q2	10.5	(4.4)	15.5	(6.0)	37.9	(5.8)
Q3	23.2	(6.9)	44.9	(9.2)	37.3	(7.6)
Q4	44.8	(8.7)	73.7	(8.1)	54.9	(8.4)
Q5	58.9	(10.1)	98.6	(2.9)	58.9	(10.1)
Net Income						
Q1	3.9	(3.0)	14.6	(5.8)	10.6	(3.4)
Q2	17.9	(5.6)	37.7	(8.5)	30.7	(6.0)
Q3	25.2	(5.9)	34.6	(6.5)	59.7	(6.9)
Q4	33.6	(9.0)	44.5	(9.7)	76.5	(6.2)
Q5	45.2	(9.2)	51.7	(9.4)	71.8	(8.2)
Canton						
Luxembourg	15.8	(3.3)	24.8	(4.2)	35.3	(4.4)
Other	19.7	(3.1)	34.4	(4.0)	38.1	(3.5)
Household size						
1	15.5	(3.4)	29.0	(4.4)	24.1	(4.0)
2	21.3	(4.8)	36.5	(6.4)	40.3	(5.6)
3	22.2	(6.7)	35.2	(7.2)	55.2	(6.9)
4	24.5	(6.8)	25.2	(6.8)	61.1	(8.2)
5+	6.9	(4.2)	21.5	(10.1)	37.6	(10.6)

Notes: The table reports the share of renting households simultaneously fulfilling purchase affordability and repayment affordability, or fulfilling each criteria separately. Standard errors are reported in parentheses. Selected household characteristics represent some of the main ingredients for the construction of the affordability measures. $Q1$ to $Q5$ denote quintiles.

Regarding the annuity factor, we take into account the mortgagee's age a_h. Two aspects are relevant here: the maximum maturity usually observed in Luxembourg is 30 years[26] and the mortgage usually needs to be completely repaid by the regular legal retirement age, i.e., upon turning 66,

$$m_h = \min\{30; \max\{0; 66 - a_h\}\}.$$

Across renting households, the median age of the reference person is 44 years. Thus, the median potential mortgage length equals 22 years. About 9% of reference persons in renting households are 66 or above, and thus cannot take out a mortgage. As a robustness check, we later vary this assumption for couples by taking into account the age of both partners.

As a final ingredient to calculate the financial margin, we need to estimate basic living costs based on survey questions.[27] The estimated mean monthly basic living costs range from EUR 814 in the lowest net income quintile to EUR 1,909 in the highest quintile.

We find a positive financial margin for 37% (2.7pp) of all current renting households. This means that they have sufficient income to cover all recurring expenses including a hypothetical mortgage financing the acquisition of their current main residence. Only 18.1% (2.2pp) of all renting households fulfil both purchase and repayment affordability ($PA \cap RA$).[28] These results reveal that home-ownership is not readily achievable for a large share of current renters in Luxembourg supporting the call by BCL Governor Gaston Reinesch (2022) that a *"coordinated policy action"* is *"key in ensuring access to home ownership for the population at large and in avoiding dividing the population into "insiders" and "outsiders""*.

Table 12 shows different categories of renting households that fulfil both purchase and the repayment affordability criteria, or either criteria in isolation. The likelihood of simultaneously fulfilling both affordability criteria decreases with age. It is also lower for households resident in the canton Luxembourg compared to the rest of the country. The likelihood of simultaneously fulfilling both affordability criteria increases with net income and net liquid assets. Single households and large families rarely fulfil both affordability criteria.

Furthermore, for 73.1% (2.7pp) of renting households, monthly interest payments (taking into account the tax deductibility of interest payments) would be lower than their reported current monthly rent paid. For the remaining 26.9% of renters, monthly interest payments would be higher than the monthly rent they reported.

To conclude, only 15.3% (2.0pp) of all renting households meet both affordability criteria to purchase their main residence, and would face lower mortgage payments than their current rent. For the vast majority, however, purchasing the dwelling they rent is not feasible at current market conditions.

A more nuanced analysis shows that in total 18.9% (2.2pp) of all renting households would earn enough income to fulfil RA yet they lack sufficient resources to meet PA. This share decreases to 17.5% (2.2pp) if we require that they would have to pay less interest than their current rent. These households thus would greatly benefit from home-ownership but are hampered by lacking wealth to overcome the down-payment constraint (PA).

5 Conclusions

Surveys routinely ask owner-occupiers to estimate the current market value of their home and ask renters to report their monthly rent. The reliability of the market valuations depends on owners' knowledge of the local housing market and their ability to apply it to their own home. Instead, reported rent is usually far from current market conditions because many tenants are on long-term rent contracts that are rarely updated in Europe's heavily regulated rental markets. Thus, survey data is a challenging source to assess housing markets.

We propose a feasible approach to incorporate market data into such surveys. For this purpose, we elicit several additional dwelling characteristics in LU-HFCS, which we match to market data via hedonic models to obtain more objective current market values for the entire residential housing stock in Luxembourg.

Not surprisingly, we find large deviations between imputed and reported rents in our technical analysis, since existing contracts tend to be slow in adjusting when conditions tighten in the market. Using imputed house price valuations instead of those reported by survey participants, we found that total wealth was significantly higher. Most households reported a value for their main residence below the imputed one. Only the most affluent households (the top 20% of the income or wealth distribution) reported a median housing value above the imputed median. As a result, measured wealth inequality appears to be lower when using imputed house price valuations.

By imputing house sales and rent prices for all dwellings in Luxembourg, our strategy also allows us to calculate price-to-rent, price-to-income and rent-to-income ratios for households in different groups and regions. We can assess the Luxembourg housing market in terms of affordability. The results suggest that 2018 price levels on Luxembourg's housing market made it impossible for the majority of renters to purchase the home they occupy given their financial situation and current housing market conditions. Given the substantial real estate price increases since 2018[29], affordability has likely worsened ever since then. If house prices rise faster than rents, the ratio of renters to owners should increase over time.

In this article we have demonstrated some potential uses of incorporating objectified sales and rent prices reflecting current market conditions into a multi-purpose country-representative survey. The resulting direct and counterfactual data provides a policy tool for micro-simulation purposes as well as population-representative housing market statistics. Such data can, however, also be used to evaluate policies targeted to support either home-ownership or renting, stress-test households' portfolios in hypothetical scenarios, or bring housing statistics forward or backwards in time using a constant housing stock in the short run and a flexible adjustment of sales and rental prices for each dwelling to market conditions (by adjusting the time window from which market data are retrieved). We thus believe that such data could greatly extend the toolkit available to policy-makers.

Notes

[1]Although rents are paid repeatedly, they usually do not reflect current market trends, as existing rent contracts are rarely adjusted. This is particularly true for rent controlled markets – the norm in Europe (see Kholodilin, 2020). Rental-equivalent methods relying on paid rents are challenged in a variety of applications (see, for instance, de Haan and Diewert, 2013; Hill et al., 2020).

[2]Syz (2008) finds that residential housing made up roughly one third of global total wealth in 2008.

[3]According to the 2011 EU Population and Housing Census, roughly 10% of the residential housing stock in Luxembourg was constructed before 1919, 32% between 1919 and 1970, 33% between 1971 and 2000 and 14% between 2001 and 2010. (For 11% the construction year is unknown.)

[4]Recently acquired homes bought-to-let represent a small and likely specific sub-segment not quite representative for the housing stock as a whole. For instance, Bracke (2015) documents differences in prices as well as dwelling characteristics in housing sales in London depending on the transaction's purpose.

[5]The HFCS is the single most important source to compile harmonised wealth-related statistics across Europe. The network is coordinated by the European Central Bank. See `https://www.ecb.europa.eu/pub/economic-research/research-networks/html/researcher_hfcn.en.html`, last accessed on September 3, 2021.

[6]See `observatoire.liser.lu`, last accessed on November 30, 2021

[7]With the best available sampling frame in Luxembourg, the Luxembourg Social Security Register, we still miss around 10% of the population including households of international civil servants and (standard for household surveys) collective households such as, for instance, homes for the elderly.

[8]A general concern could be the potentially difficult task to report the monthly rent *excluding* charges.

[9]According to the Minister of Economics, in 2019 there were 1,221 registered real estate agencies in Luxembourg and the number was sharply increasing over the last years (see Luxembourg Times, November 24, 2019, https://www.luxtimes.lu/en/luxembourg/real-estate-firms-flood-into-luxembourg-as-prices-soar-619e68e7de135b9236549a12).

[10]See the Luxembourg Data Platform for population sizes per postcode: `https://data.public.lu/en/datasets/population-par-code-postal-population-per-postal-code` (last accessed on January 11, 2021) and STATEC for de jure mid-year total population sizes: 596,336 (2017), 607,950 (2018), 620,001 (2019).

[11]Currently, precise data on average listing times for dwellings do not exist in Luxembourg. Data on "days on market" used for measuring tightness however exist, e.g., in the US https://fred.stlouisfed.org/series/MEDDAYONMARUS (last accessed on November 29, 2021).

[12]The capital Luxembourg City has a population of nearly five times that of the second largest urban area around Esch-sur-Alzette, which in turn is 40% larger than the third largest town. For further information, see the latest population per municipality figures provided by STATEC: `https://statistiques.public.lu/stat/TableViewer/tableView.aspx`, last accessed on January 11, 2021

[13]The following interviewer ratings did not enter the final specification as coefficients were not significant. These are understanding questions, reliability of income and wealth information, ability to express amounts in EUR, ease in responding, ability to express himself/herself, and three ratings of the dwelling, namely the outward appearance, the comparison to the neighbourhood and the rating of surrounding buildings.

[14] Details are summarised here: `http://www.statistiques.public.lu/stat/TableViewer/document.aspx?ReportId=13442&IF_Language=eng&MainTheme=4&FldrName=4&RFPath=35` (Last accessed October 21, 2021).

[15]The harmonised HFCS core questionnaire only asks for gross household income. For Luxembourg, the Pearson correlation coefficient between gross and net household income is $\rho \approx 0.8$.

[16]See housing price indicators provided by the OECD, `https://doi.org/10.1787/54a3bf57-en` (Last accessed September 8, 2021).

[17]See `https://guichet.public.lu/en/citoyens/logement/acquisition/aides-capital/prime-construction-acquisition.html` (Last accessed November 8, 2021) for details.

[18]See `http://data.legilux.public.lu/file/eli-etat-leg-rcsf-2020-12-03-a969-jo-fr-pdf.pdf` (Last accessed November 9, 2021).

[19] Net liquid assets exclude typically illiquid assets such as other non-mortgage loans and the value of non-self-employment private businesses.

[20] See `https://www.globalpropertyguide.com/Europe/Luxembourg/Buying-Guide` (Last accessed November 9, 2021)

[21] Standard errors are reported in brackets.

[22] Source: ECB Statistical Data Warehouse, `https://sdw.ecb.europa.eu/quickview.do?SERIES_KEY=124.MIR.M.LU.B.A2C.A.R.A.2250.EUR.N` (Last accessed November 9, 2021).
Our model does not simulate the impact of the “Subvention et bonification d’intérêt” (see LISER, 2022), which reduces the mortgage rate of a household for purchasing a home. Eligibility and size depend on the household’s income and family situation. At the same time, low household income requires a risk premium on the mortgage rate granted by credit institutions. We argue that the inverse relationship between “Subvention et bonification d’intérêt” and the risk premium on mortgage rates is mutually offsetting and therefore exclude both factors from our simulation.

[23] See `https://guichet.public.lu/en/citoyens/logement/acquisition/aides-indirectes/declarer-residence-principale-secondaire.html` (Last accessed November 8, 2021)

[24] The maximum amount decreases step-wise for subsequent years.

[25] This assumption is less strong as it may seem at first sight. In tax class 1 (single taxpayer), a marginal tax rate of 39% is charged between EUR 46,700 and EUR 100,750 of taxable income. This constitutes the most common marginal tax bracket. See also an analysis of the Luxembourg Aconseil Economique et Social, `https://ces.public.lu/dam-assets/fr/avis/prix-salaires/2015-fiscalite.pdf` (Last accessed November 8, 2021).

[26] The 2019 law, see `http://data.legilux.public.lu/eli/etat/leg/loi/2019/12/04/a811/jo`, allows the *Commission de Surveillance du Secteur Financier* (CSSF) to set a maximum limit for the original loan maturity between 20 and 35 years. Yet, LU-HFCS data show that 99% of mortgages for the acquisition of the main residence have an initial maturity of at most 30 years. Bank statistics on the outstanding stock of mortgages to the household sector (`https://www.bcl.lu/fr/statistiques/series_statistiques_luxembourg/11_etablissements_credit/11_07_Tableau.xlsx`, Table 11.07 – version 04/30/2021) report that the share of mortgages with an initial maturity of more than 30 years is 7.4% (average across 2018). The difference between bank statistics and HFCS results is likely due to fundamentally different weighing concepts. The HFCS attributes weights to households to gain a population-representative total with regard to socio-economic characteristics while bank mortgage statistics are weighted by the outstanding amounts. Both bank and HFCS statistics support that a maximum maturity of 30 years is a reasonable assumption for our analysis. (Hyperlinks last accessed November 8, 2021)

[27] We define monthly costs as the sum of food consumption at home, 50% of food consumption outside home, amount spent on utilities and 10% of a household’s net equivalised income (using the square root equivalence scale, see Atkinson et al., 1995; OECD, 2015) to approximate a set of minimum expenditures not covered by the items before (e.g., for clothing, health care or mobility). All these ingredients are available in the LU-HFCS and thus reflect individual realised spending behaviour of the household assumed to purchase the home.

[28] These results are calibrated using the age of the reference person to determine the maximum maturity for a specific household. By that, we measure an average age of 45.7 years. As it is common for couples to jointly purchase homes, both partners’ ages may be relevant in practice. Thus, we estimate an upper and lower bound for the share of households fulfilling RA and $PA \cap RA$, respectively. We re-estimate results using the couple’s minimum (average 44.2 years) and maximum (average 46.5 years) age yielding an interval of plausible results. This translates for RA into $(34.8\%; 39.5\%)$ and for $PA \cap RA$ into $(17.34\%; 18.75\%)$. By construction, the resulting intervals overlap the point estimates and are rather narrow. Thus, we proceed with the age of the reference person as main result.

[29] Real estate prices for new and existing dwellings increased almost 13% per year on average from 2018Q3 to 2021Q3, see `https://statistiques.public.lu/stat/tableviewer/document.aspx?ReportId=13440`.

References

Agarwal, S. (2007). The impact of homeowners' housing wealth misestimation on consumption and saving decisions. *Real Estate Economics*, 35(2):135–154.

Alexeev, S. (2020). The role of imputed rents in intergenerational income mobility in three countries. *Journal of Housing Economics*, 49:101710.

Atkinson, A. B., Rainwater, L., and Smeeding, T. M. (1995). *Income Distribution in OECD Countries: Evidence from the Luxembourg Income Study*. OECD, Paris.

Bach, S., Thiemann, A., and Zucco, A. (2019). Looking for the missing rich: Tracing the top tail of the wealth distribution. *International Tax and Public Finance*, 26(6):1234–1258.

Benítez-Silva, H., Eren, S., Heiland, F., and Jiménez-Martín, S. (2015). How well do individuals predict the selling prices of their homes? *Journal of Housing Economics*, 29:12–25.

Bracke, P. (2015). House prices and rents: Microevidence from a matched data set in Central London. *Real Estate Economics*, 43(2):403–431.

Chen, Y., Mathä, T. Y., Pulina, G., Schuster, B., and Ziegelmeyer, M. (2020). The Luxembourg Household Finance and Consumption Survey: Results from the third wave. *BCL Working Papers*, 142.

Cook, R. D. (1977). Detection of influential observation in linear regression. *Technometrics*, 19(1):15–18.

de Haan, J. and Diewert, W. E., editors (2013). *Handbook on Residential Property Prices Indices (RPPIs)*. Methodologies and Working Papers. Eurostat, Luxembourg.

EG-LMM (2020). Understanding household wealth: linking macro and micro data to produce distributional financial accounts. *Statistics Paper Series*, No 37 / July 2020.

ESA (2010). *European System of Accounts (ESA 2010)*. Eurostat, European Commission, Luxembourg.

Gan, Q. and Hill, R. J. (2009). Measuring housing affordability: Looking beyond the median. *Journal of Housing Economics*, 18(2):115–125.

Garner, T. I. and Verbrugge, R. (2009). Reconciling user costs and rental equivalence: Evidence from the US consumer expenditure survey. *Journal of Housing Economics*, 18(3):172–192.

Girshina, A., Koulischer, F., and von Lilienfeld-Toal, U. (2021). Housing affordability and transaction tax subsidies. *Available at SSRN 3758466*.

Glaesener, M.-L. and Caruso, G. (2015). Neighborhood green and services diversity effects on land prices: Evidence from a multilevel hedonic analysis in Luxembourg. *Landscape and Urban Planning*, 143:100–111.

Goodman Jr, J. L. and Ittner, J. B. (1992). The accuracy of home owners' estimates of house value. *Journal of Housing Economics*, 2(4):339–357.

Hicks, J. R. et al. (1975). Value and capital: An inquiry into some fundamental principles of economic theory. *OUP Catalogue*.

Hill, R. J., Steurer, M., and Waltl, S. R. (2020). Owner-occupied housing, inflation and monetary policy. *Graz Economics Papers*, 2020–18.

Hill, R. J. and Syed, I. A. (2016). Hedonic price-rent ratios, user cost, and departures from equilibrium in the housing market. *Regional Science and Urban Economics*, 56:60–72.

Himmelberg, C., Mayer, C., and Sinai, T. (2005). Assessing high house prices: Bubbles, fundamentals and misperceptions. *Journal of Economic Perspectives*, 19(4):67–92.

Household Finance and Consumption Network (2020). HFCS user database: Documentation core and derived variables. *HFCS*, March-2020.

Kaempff, B. (2018). Les interventions de l'etat sur le marche immobilier au Luxembourg. *BCL Bulletin*, 2018-01.

Kain, J. F. and Quigley, J. M. (1972). Note on owner's estimate of housing value. *Journal of the American Statistical Association*, 67(340):803–806.

Kennickell, A. B. (2019). The tail that wags: differences in effective right tail coverage and estimates of wealth inequality. *The Journal of Economic Inequality*, 17(4):443–459.

Kholodilin, K. (2020). Long-term, multicountry perspective on rental market regulations. *Housing Policy Debate*, 30(6):994–1015.

Kiel, K. A. and Zabel, J. E. (1997). Evaluating the usefulness of the American Housing Survey for creating house price indices. *The Journal of Real Estate Finance and Economics*, 14(1-2):189–202.

Kiel, K. A. and Zabel, J. E. (1999). The accuracy of owner-provided house values: The 1978–1991 American Housing Survey. *Real Estate Economics*, 27(2):263–298.

Kish, L. and Lansing, J. B. (1954). Response errors in estimating the value of homes. *Journal of the American Statistical Association*, 49(267):520–538.

Kolbe, J., Schulz, R., Wersing, M., and Werwatz, A. (2020). Real estate listings and their usefulness for hedonic regressions. *Empirical Economics*, 61:3239–3269.

Lepinteur, A. and Waltl, S. R. (2021). Tracking owners' sentiments: Subjective home values, expectations and house price dynamics. *LISER Working Papers*, 2021-02.

Linneman, P. and Wachter, S. (1989). The impacts of borrowing constraints on homeownership. *Real Estate Economics*, 17(4):389–402.

LISER (2022). L'impact des politiques sociales et fiscales en matière de logement sur la situation de revenu des locataires et propriétaires. *Observatoire de l'habitat*, La Note 30:1–62.

Mathä, T. Y., Porpiglia, A., and Ziegelmeyer, M. (2017). Household wealth in the euro area: The importance of intergenerational transfers, homeownership and house price dynamics. *Journal of Housing Economics*, (35):1–12.

McFadden, D. (1974). *Frontiers in Econometrics*, chapter Conditional logit analysis of qualitative choice behavior, pages 105–142. Academic Press, New York.

Molloy, R. and Nielsen, E. R. (2018). How can we measure the value of a home? comparing model-based estimates with owner-occupant estimates. *FEDS Notes*, (2018-10):11.

OECD (2015). What are equivalence scales? *OECD Project on Income Distribution and Poverty*, Available at: www.oecd.org/els/soc/OECD-Note-EquivalenceScales.pdf.

OECD (2021). Building for a better tomorrow: Policies to make housing more affordable. *Employment, Labour and Social Affairs Policy Briefs*, Available at: http://oe.cd/affordable-housing-2021.

Reinesch, G. (2022). Residential real estate prices in Luxembourg. *Blog post by Gaston Reinesch, Governor of the BCL*, https://www.bcl.lu/en/publications/Blog/Blog-14/index.html.

Rosen, H. S., Rosen, K. T., and Holtz-Eakin, D. (1984). Housing tenure, uncertainty, and taxation. *The Review of Economics and Statistics*, 66(3):405–416.

Rosen, S. (1974). Hedonic prices and implicit markets: Product differentiation in pure competition. *Journal of Political Economy*, 82(1):34.

Silverman, B. W. (1986). *Density Estimation for Statistics and Data Analysis.* Chapman and Hall, London; New York.

Syz, J. M. (2008). *Property Derivatives: Pricing, Hedging and Applications.* John Wiley & Sons, Chichester, England.

Vermeulen, P. (2018). How fat is the top tail of the wealth distribution? *Review of Income and Wealth*, 64(2):357–387.

Waltl, S. R. (2016). A hedonic house price index in continuous time. *International Journal of Housing Markets and Analysis*, 9(4):648–670.

Waltl, S. R. (2018). Estimating quantile-specific rental yields for residential housing in Sydney. *Regional Science and Urban Economics*, 68:204–225.

Waltl, S. R. (2022). Wealth inequality: A hybrid approach toward Multidimensional Distributional National Accounts in Europe. *Review of Income and Wealth*, 68(1):74–108.

Waltl, S. R. and Chakraborty, R. (2022). Missing the wealthy in the HFCS: Micro problems with macro implications. *Journal of Economic Inequality*, https://doi.org/10.1007/s10888-021-09519-1.

Wood, S. N. (2006). *Generalized additive models: an introduction with R.* Chapman and Hall / CRC, Boca Raton, Florida.

Appendix

A Imputation Strategy

A.1 External Data Sources

The imputation models rely on a wide range of external data. In particular, we employ data on advertised units to sell or rent, as well as notary deed data collected by the Housing Observatory, covering a period of nearly a decade,[30] information on housing stocks per municipality from the housing census,[31] geographic information on municipal[32] and neighbourhood[33] (in the case of the two large urban agglomerations of Luxembourg City and Esch-sur-Alzette) boundaries, as well as data on postal code centroids published on the Luxembourg Open Data portal.[34]

Additionally to rough location fixed-effects, we use a measure of reachability. Following Glaesener and Caruso (2015), we use the distance to Luxembourg's capital as an alternative predictor in the hedonic models.

A.2 Linking Data Sets

We create a imputation model estimated on the basis of advertised housing units for sale and for rent. To ensure comparability, they relate to the same year as the fieldwork of the LU-HFCS (i.e., 2018). We use characteristics reported in advertisements linkable to characteristics we generically available or specifically included in the LU-HFCS. These are: dwelling type, dwelling surface, plot surface, year of construction, number of bedrooms, energy score, dwelling state and location.

In a data cleaning process, we exclude re-advertisements of the same property, leaving us with 14,759 advertisements of properties for sale and 7,860 for rent in 2018.

Real estate advertisements do not consistently include all the characteristics mentioned above as only few pieces of information are required by advertising portals. We address this shortcoming through the following assumptions and categorisations. Missing information on the construction year affects 11,459 entries in the sales set, and 4,546 entries in the rents set. In such cases, we fill gaps with an average municipality-specific building age to proxy the place-specific urban development regime. We compute this average from STATEC census data reporting the distribution of building age per municipality. In order to reduce insecurity, we subsequently group building years into six brackets, informed by the age distribution of advertised property in the wider Housing Observatory data set: pre-1950, 1950-1969, 1970-1989, 1990-1999, 2000-2009 and 2010–2020.

Table 13 – Correspondence Table – Housing Observatory vs. HFCS.

	HFCS					Housing Observatory	
	Question	Code	Reported by	Original Scale	Trans-formation	Original Scale	Trans-formation
Type	What kind of accommodation is it?	hlb0100	R	1=detached house, 2=semi-detached house, 3=terraced/ row/town house, 4=a farm, 5=apartment/ flat in a building of 2 to 4 acc., 6= apartment/flat in a building of 5 to 9 acc., 7=apartment/flat in a building with 10 or more acc., 8=another type of acc.	1–4 = house, 5–7 = flat	house or flat	–
Surface	What is the size of the residence in sqm?	hlb0100	R	Numeric value	–	Numeric value	–
Plot size	In ares or square meters, how big is the land/construction site belonging to your household main residence?	hlb0200x	R	Numeric value	–	Numeric value	–

Construction	In which year was the residence originally constructed; more specially, when was the construction completed?	hlb0210	R	Year	0=pre-1950, 1=1950-1969, 2=1970-1989, 3=1990-1999, 4=2000-2009, 5=post-2010	Year	0=pre-1950, 1=1950-1969, 2=1970-1989, 3=1990-1999, 4=2000-2009, 5=post-2010
Bedrooms	What is the number of bedrooms in the residence?	hlb0310	R	Numeric value	1=1 room, 2=2 rooms, 3=3 rooms, 4=4+ rooms	Numeric value	1=1 room, 2=2 rooms, 3=3 rooms, 4=4+ rooms
Energy Score	What is the energy performance class of the residence?	hlb0410	R	Classes A–I	0=A-D, 1=E-G, 2=H-I	Classes A–I	0=A-D, 1=E-G, 2=H-I
Status	Specify the outward appearance of the dwelling/building.	sc0400	I	1 = Generally clean and sound, 2 = Some peeling paint or cracks in walls, 3 = Needs substantial painting, refilling or repair, 4 = Dilapidated	1-2 = 4 Good, 3-4 = 7 Average	1 = New-build, 4 = Good, 7 = Average	–
Location	What is the post code where the dwelling is located?	hl0100	R	Post code	Conversion to composite geography	Municipality and neighbourhood	Conversion to composite geography

Notes: HFCS item codes follow the EU-wide harmonised coding scheme whenever applicable and Luxembourg-specific codes otherwise (see Household Finance and Consumption Network, 2020). We distinguish between information provided by the interviewee (R) during the survey interview and assessments provided by the interviewer (I).

We implement a similar strategy for filling in gaps in the data on the energy class of the dwelling, which affects 4,653 cases in the sales data set and 4,360 in the rents data set. Based on the insights into energy class distributions across municipalities and building periods gained from the larger Housing Observatory data set, we group energy performance classes into three brackets: high (A-D), medium (E-G) and low (H-I). We then study their distribution across municipalities and the building periods described above, subsequently replacing missing values with median values per building bracket and municipality from the wider data set. For homes built during periods for which we have no data for that municipality in the wider Housing Observatory data set, we use the energy class values corresponding to the closest period for which data is available in that municipality.

We remove observations where the number of rooms or the dwelling surface are missing, as this concerns less than 1,500 cases in the combined sales and rents data set. Although it could be argued to re-construct the number of rooms from observed cases, prior research suggests that there are significant differences in dwelling surface between and within municipalities, which could result in further bias.[35] We also group dwellings with four or more rooms into a generic 4+ category following a standard diminishing marginal rate of return argument. Plot surface was assumed to be 0 for adverts that did not mention it (most correspond to flats, where the likelihood of a garden is low), while the dwelling state was marked as average for advertisements that did not mention any improvements or the need for any repairs.

On the location side, the data from the Housing Observatory is reported at the communal and village/neighbourhood level. While communal levels are sufficiently fine-grain for most locations in Luxembourg (featuring population sizes amounting from just under 1,000 to around 25,000[36]), two urban areas feature higher population and a larger number of LU-HFCS respondents: Luxembourg City (with a population of over 120,000) and Esch-sur-Alzette (just over 36,000). In order to account for this difference, we build a composite geographical level by splitting Luxembourg City into its composing neighbourhoods, separating Esch-sur-Alzette and its largest suburb Belval, and preserving the communal level for all other locations. The advantage of this strategy is that it preserves the link to administrative levels, the key to sourcing complementary information and to linking the model based on the Housing Observatory data to the HFCS survey.

A.3 Imputation Models

We estimate the hedonic model (1) relying on all linkable predictors (re-)coded as described in Table 13.

The model's performance is slightly improved (increase in adjusted R^2 by one percentage point) with the exclusion of five influential statistical outliers from the sales data set identified via Cook's distance (see Cook, 1977). Table 14 reports estimation results.

As we impute prices for dwellings per construction not part of the training sample, we estimate the potential imputation error from an out-of sample prediction check. Therefore, we leave out 20% of the data used for validation. These 20% are chosen via geographically stratified random sampling.

The selected hedonic model is subsequently used to impute sales and rent prices for each dwelling described in the HFCS survey. For this article, we added several questions describing the physical features of the main residence as well as a question about its location. For questions where the respondents refused to answer, five multiply imputed implicates are introduced. For apartments no plot size is reported. The following physical dwelling characteristics elicited in the HFCS[37] are used to impute prices, after ensuring that they are transformed into the same

Table 14 – Hedonic Regression Results

	log(P^S)					log(P^R)				
	(S.Main)	(S.Alt.1)	(S.Alt.2)	(S.Alt.3)	(S.Alt.4)	(R.Main)	(R.Alt.1)	(R.Alt.2)	(R.Alt.3)	(R.Alt.4)
Intercept	12.104***	12.469***	12.25***	12.347***	12.446***	6.276***	6.834***	12.252***	6.456***	6.684***
	(0.037)	(0.020)	(0.016)	(0.037)	(0.019)	(0.160)	(0.038)	(0.016)	(0.158)	(0.032)
Dwelling characteristics.										
Type of Dwelling										
Flat					*(ref. cat.)*					
House	0.493***	0.543***	0.488***	0.031***	-0.103***	0.265***	0.310***	0.488***	0.239***	0.082***
	(0.014)	(0.018)	(0.014)	(0.006)	(0.009)	(0.025)	(0.027)	(0.014)	(0.012)	(0.016)
Surface	0.006***	0.007***	0.006***	0.003***	0.003***	0.004***	0.004***	0.006***	0.004***	0.004***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Plot size	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Construction year										
pre 1950					*(ref. cat.)*					
1950-1969	0.089***	0.134***	0.093***	0.066***	0.051**	0.228***	0.120**	0.093***	0.042 .	-0.020
	(0.014)	(0.018)	(0.014)	(0.010)	(0.016)	(0.036)	(0.037)	(0.014)	(0.022)	(0.029)
1970-1989	0.118***	0.229***	0.119***	0.078***	0.223***	0.2035***	0.1091*	0.119***	0.028	-0.0641*
	(0.022)	(0.028)	(0.022)	(0.012)	(0.018)	(0.043)	(0.045)	(0.022)	(0.024)	(0.031)
1990-1999	0.162***	0.192***	0.164***	0.118***	0.088***	0.266**	0.0380	0.164***	0.037	-0.145***
	(0.038)	(0.049)	(0.038)	(0.013)	(0.020)	(0.099)	(0.107)	(0.038)	(0.026)	(0.034)
2000-2009	0.146*	0.196*	0.140 .	0.113***	0.102***	0.292*	0.155	0.140 .	0.054*	-0.111***
	(0.073)	(0.094)	(0.073)	(0.012)	(0.019)	(0.119)	(0.128)	(0.073)	(0.024)	(0.031)
2010-2020	0.216	0.330	0.205	0.040***	0.028	0.296**	0.288*	0.205	0.078**	-0.024
	(0.203)	(0.262)	(0.204)	(0.011)	(0.018)	(0.105)	(0.116)	(0.204)	(0.024)	(0.032)
Number of bedrooms										
1					*(ref. cat.)*					
2	0.143***	0.100***	0.143***	0.242***	0.181***	0.119***	0.112***	0.143***	0.120***	0.092***
	(0.006)	(0.008)	(0.006)	(0.005)	(0.009)	(0.006)	(0.007)	(0.006)	(0.006)	(0.008)
3	0.182***	0.122***	0.183***	0.375***	0.320***	0.165***	0.146***	0.183***	0.166***	0.154***
	(0.008)	(0.011)	(0.008)	(0.006)	(0.010)	(0.010)	(0.011)	(0.008)	(0.009)	(0.013)
4+	0.226***	0.150***	0.224***	0.419***	0.372***	-0.137***	-0.135***	0.224***	-0.145***	-0.109***
	(0.010)	(0.013)	(0.010)	(0.008)	(0.013)	(0.009)	(0.010)	(0.010)	(0.009)	(0.012)
Energy Efficiency Class										
low (H-I)					*(ref. cat.)*					
medium (E-G)	0.063***	0.083***	0.068***	0.032***	0.155***	0.317***	0.250***	0.068***	0.032**	0.243***
	(0.019)	(0.024)	(0.018)	(0.006)	(0.010)	(0.045)	(0.048)	(0.018)	(0.010)	(0.012)
high (A-D)	0.157*	0.268***	0.141*	0.132***	0.217***	0.358***	0.285*	0.141*	0.123***	0.290***
	(0.061)	(0.077)	(0.060)	(0.007)	(0.011)	(0.106)	(0.116)	(0.060)	(0.013)	(0.016)
Dwelling State										
planned					*(ref. cat.)*					
new-build	0.009	-0.043*	0.003	0.016	-0.021	-0.006	-0.029	-0.005	0.004	0.003
	(0.014)	(0.018)	(0.014)	(0.015)	(0.024)	(0.019)	(0.022)	(0.014)	(0.020)	(0.026)
good	-0.032***	-0.085***	-0.038***	-0.044***	-0.065***	-0.033**	-0.020 .	-0.038***	-0.039***	0.040**
	(0.007)	(0.009)	(0.007)	(0.007)	(0.011)	(0.010)	(0.011)	(0.007)	(0.010)	(0.013)
average	-0.054***	-0.078***	-0.059***	-0.058***	-0.073***	-0.059***	-0.052***	-0.059***	-0.063***	-0.038***
	(0.005)	(0.006)	(0.005)	(0.005)	(0.008)	(0.008)	(0.009)	(0.005)	(0.008)	(0.011)
Interactions										
Dwelling Type × Surface	-0.003***	-0.004***	-0.003***	✗	✗	0.000	0.000	-0.003***	✗	✗
	(0.000)	(0.000)	(0.000)			(0.000)	(0.000)	(0.000)		
Construction Year × Energy Class	✓	✓	✓	✗	✗	✓	✓	✓	✗	✗
Geographical characteristics.										
Locality Dummies	✓	✗	✗	✓	✗	✓	✗	✗	✓	
Distance to Capital	✗	-0.014***	✗	✗	✗	✗	-0.013***	✗	✗	✗
		(0.000)					(0.000)			
Linked neighbour Spline	✗	✗	✓	✗	✗	✗	✗	✓	✗	✗

Notes: P^S and P^R denotes the advertised sales or monthly rent price, respectively. Coefficients for single localities are left out. We indicate in- and exclusion of a (set) of variables by ✓and ✗, respectively. Significance is indicated using standard notation: $^{\cdot}$ p-value<0.1; *p-value<0.05; **p-value<0.01; ***p-value<0.001.
Source: Authors' calculations based on data from the *Housing Observatory*, on advertisements available between 1 January and 31 December 2018.

scales/categories used for the Housing Observatory model: dwelling type (hlb0100), dwelling surface (hb0100), plot surface (hlb0200x), year of construction (hlb0210), number of bedrooms (hlb0310), energy score (hlb0410) and dwelling status (sc0400). See Table 13 for details.

As locational measure, we elicit a postcodes (hl0100) in the course of the LU-HFCS survey interview. Luxembourg features an exceptionally fine grained postcode system: There are 4,022 postcodes, with a maximum population per postcode of 1,500. On average, there are only 158 persons per code.[38] Yet, the codes have the distinct disadvantage of occasionally crossing administrative boundaries at neighbourhood and municipal levels. We therefore convert the post codes into our composite geography by applying a juxtaposition strategy in QGIS: the geographical centroids of postcodes are superimposed to the boundaries of our composite geography. For postcodes that fall across municipal lines, the centroids cover municipal sections of the shared postcode; consequently, the resulting enhanced attribute table allocates multiple municipalities to single postcodes that cover residences in different administrative units.

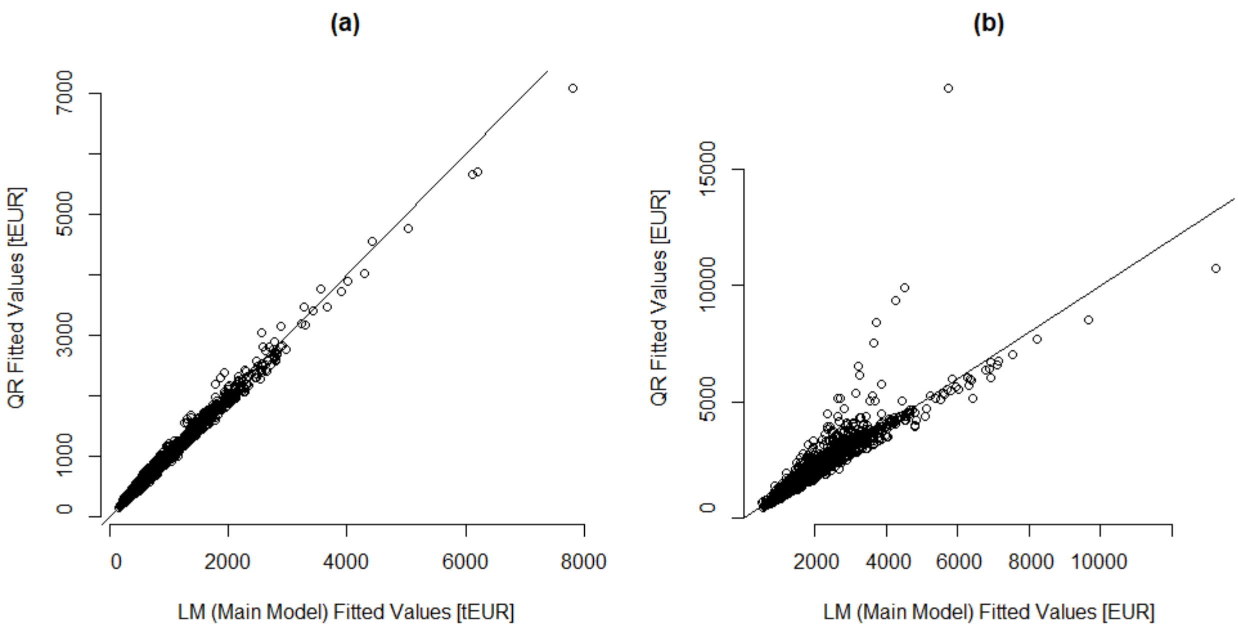
This non-unique allocation is acknowledged in the imputation exercise through the introduction of geographical implicates for LU-HFCS respondents located in postcodes falling across municipal lines. In order to ensure no systematic bias, we randomise the municipal allocations across implicates.

A.4 Robustness Analyses

In subsection 3.1, we perform several robustness and sensitivity analyses as part of the model selection procedure and conclude that the models S.Main and R.Main are best performing along several dimensions. Here, we cross-check predictions using the same specification but change model classes: we switch from linear to quantile regression models evaluated at the median. Both model classes yield unbiased predictions for median rent or sale prices conditional on housing characteristics.

We call the resulting quantile models for the median QR.Sale and QR.Rent. They perform similarly as their linear counterparts in terms of goodness-of-fit: AIC -7,276 (QR.Sale) and -4,204 (QR.Rent); BIC -6,105 (QR.Sale) and -3,232 (QR.Rent). Reassuringly, imputations from either model also strongly correlate as shown in Figure O.1. This is particularly true for sales prices.

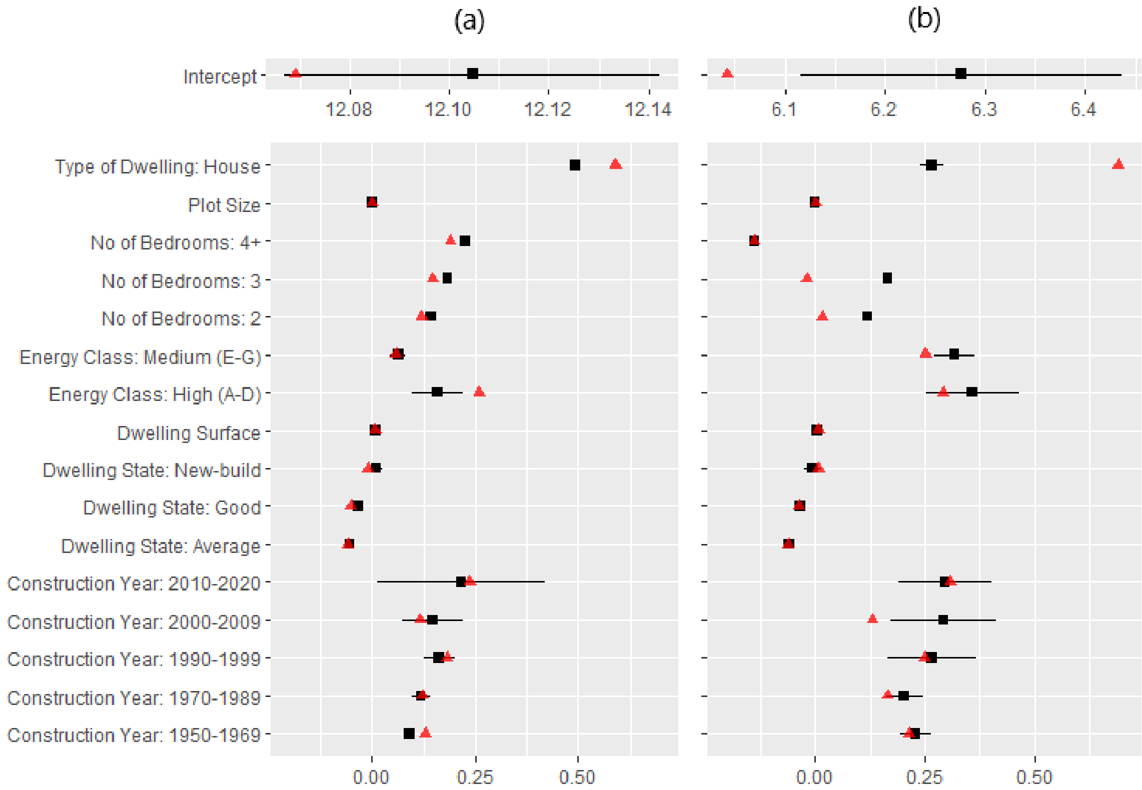
Figure 8 – Comparison of Linear and Quantile Model Fitted Values



Notes: Panel (a) relates to imputed sales and panel (b) to imputed rent prices. Imputations use the main model specification targeted to produce median estimators but differ in model class: linear model versus quantile regression.

Diving into details, we make use of the fact that the imputations exclusively rely on the estimated model parameters and, thus, further compare estimated parameters between the main models and their exact same quantile regression counterparts. For imputing market prices, we just use estimated coefficients. Thus, we check for each coefficient, whether they differ significantly between the best linear and quantile model. We therefore construct symmetric 95%

Figure 9 – Comparison of Linear and Quantile Model Coefficients



Notes: Panel (a) relates to the sales models and panel (b) to the rent models. The black square and line represent the Linear Model coefficient and CI respectively, while the red triangle represents the Quantile Model coefficient. The linear and quantile models also contain interaction terms as well as locality dummies, but these are suppressed here.

confidence intervals for each estimated parameter and check whether this interval overlaps the corresponding estimated parameter from the QR model.

Figure 9 reports the results. The vast majority of coefficients are not statistically significantly different. Thus, either model produces comparable imputation results.

Online Appendix: Supplemental Materials

This appendix provides additional back-up information, tables and figures supporting the results discussed in the main text.

O.1 Geographical Variation

Table O.1 – Regional Variation in Prices, Rents and Household Income

		All tenure stati	Owner-occupiers	Renters
Capellen	Sales Price	767,704	818,268	543,656
	Rent Price	1,789	1,887	1,381
	Net Income	56,747	59,146	30,794
Clervaux & Vianden	Sales Price	516,755	548,462	391,806
	Rent Price	1,170	1,245	1,014
	Net Income	44,947	50,651	40,475
Diekirch	Sales Price	568,089	646,221	414,462
	Rent Price	1,286	1,408	1,083
	Net Income	52,600	69,658	39,932
Echternach	Sales Price	595,845	638,001	379,219
	Rent Price	1,462	1,523	1,165
	Net Income	63,478	65,440	31,737
Esch sur Alzette	Sales Price	573,350	623,117	428,415
	Rent Price	1,521	1,657	1,278
	Net Income	47,767	55,960	33,006
Grevenmacher	Sales Price	692,853	730,503	479,315
	Rent Price	1,616	1,750	1,354
	Net Income	58,420	59,000	52,500
Luxembourg	Sales Price	803,947	980,994	689,166
	Rent Price	1,863	2,087	1,682
	Net Income	54,429	69,727	43,084
Mersch	Sales Price	667,878	766,323	532,555
	Rent Price	1,639	1,765	1,327
	Net Income	58,880	66,640	31,094
Redange & Wiltz	Sales Price	549,545	563,862	409,975
	Rent Price	1,394	1,433	1,269
	Net Income	55,816	56,616	43,878
Remich	Sales Price	710,181	788,759	412,993
	Rent Price	1,569	1,856	1,269
	Net Income	41,584	48,068	32,488
Complete sample	Sales Price	654,104	696,994	516,697
	Rent Price	1,622	1,744	1,411
	Net Income	51,475	60,000	36,323

Notes: The table reports imputed median sales and rent price as well as median household income by canton. Smaller cantons are merged.

Figure O.1 – Residuals by Composite Geography and Canton

Figure O.1 – Residuals by Composite Geography and Canton (cont'd)

Notes: The figure extends Figure 3 reporting results by composite geography and canton. Panel (*a*) refers to sales and panel (*b*) to rents.

Source: LU-HFCS, 3rd wave.

O.2 Elicitation of (hypothetical) rent and sales prices

In the section collecting details about the household main residence, owners and renters beliefs regarding the (hypothetical) rent and sales pieces are elicited via the questions printed below. We report questions in English, however, participants could choose the language of the interview (English, French or German). Interviewers were trained to offer further clarifications in Luxembourgish or Portuguese with the help of a glossary explaining financial terms.

Owner-reported sales price – ORS. What is the value of this property, i.e. if you could sell it now how much do you think would be the price of it?

Owner-reported rent price – ORR. If you had to rent at the market price, how much do you think you would pay for your accommodation on a monthly basis? Please exclude utilities, heating, etc. if possible.

Renter-reported sales price – RRS. What do you think is the value of your residence, i.e. if you would buy it now how much do you think you would have to pay for it?

Renter-paid rent price – RPR. What is the monthly amount paid as rent (please exclude utilities, heating, etc. if possible)?

Survey participants are primarily asked to provide a numeric answer expressed in EUR. And for these four questions indeed the vast majority provide such an exact amount.[39] If respondents do not provide exact values to these four questions, respondents are asked to provide an upper or lower bound. Alternatively, a respondent is shown a scale with brackets and she is asked to indicate the likely bracket the value falls into.[40] Only if a respondent refuses to select such a bracket, the answer remains blank and will be imputed *ex post*. Missing values are multiply imputed taking upper and lower bounds into account if provided by respondents (see Chen et al., 2020). The imputation of the LU-HFCS data set is different to the one presented here in that it predominantly aims to retain internal consistency of reported or imputed values. Usually, no large-scale application of external data is used.

O.3 Definitions of Non-standard Variables

We mainly rely on survey questions taken for the Europe-wide harmonised catalogue (see Household Finance and Consumption Network, 2020). Just like in most countries conducting the HFCS survey, the harmonised questionnaire is amended by specific questions only feasible and/or relevant for a specific country. Non-standard questions we added to the questionnaire are described here.

The LU-HFCS was conducted exclusively via personal interviews that almost always took place at the reference person's home, which is also the central study object in this article. Thus, next to information reported directly by survey respondent (R), we can rely on additional assessments provided by the interviewer (I).

Throughout the article, we make use of the following country-specific and interviewer questions. We report the original scale, i.e., all options an interviewer could choose. Adjustments to the scale for harmonisation and sparsity reasons are clarified in Table 13.

Size of plot of the main residence (HLB0200x)
Reported by: Respondent.
Instructions: "In ares or square meters, how big is the land/construction site belonging to your household main residence?"

Numerical value in square meters
a set of 2 variables:
a - Surface in ares (1 are = 100 square meters)
b - Surface in square meters.

Year of construction of the main residence (HLB0210)
Reported by: Respondent.
Instructions: "In which year was the residence originally constructed; more specially, when was the construction completed?"
Numerical value

Number of bedrooms in the main residence (HLB0310)
Reported by: Respondent.
Instructions: "What is the number of bedrooms in the residence?"
Numerical value

Energy class of the main residence (HLB0410)
Reported by: Respondent.
Instructions: "What is the energy performance class of the residence ('classe de performance énergétique')?"
1 – A
2 – B
3 – C
4 – D
5 – E
6 – F
7 – G
8 – H
9 – I

Total net income (HLG0100)
Reported by: Respondent.
Instructions: "Now considering the sum of all sources of income, you said that the total gross income of your household in 2017 was <TOTAL AMOUNT OF GROSS INCOME>. How much do you think is your annual household NET income?"
Numerical value in EUR

Exterior conditions (SC0200)
Reported by: Interviewer.
Instructions: "Dwelling rating"
1 - Luxury
2 - Upscale
3 - Mid-range
4 - Modest
5 - Low-income

Outward appearance (SC0400)

Reported by: Interviewer.
Instructions: "Dwelling - outward appearance"
1 - Generally clean and sound
2 - Some peeling paint or cracks in walls
3 - Needs substantial painting, refilling or repair
4 - Dilapidated

Rating: Interior conditions (HR0200)
Reported by: Interviewer.
Instructions: "Could you describe the conditions of the interior of the dwelling?"
1 - Excellent. Walls and ceilings have no cracks, paint of panelling in good condition.
2 - Good. Needs some minor painting or refinishing.
3 - Fair. Needs major interior work. Holes and/or cracks need patching. Painting needed. etc.
4 - Poor. Some walls or ceilings need replacement.
5 - Interviewer has not seen/visited the dwelling

O.4 Additional Tables

Table O.2 – Summary Statistics: Dwelling Characteristics – HFCS

	Surface [m^2]	Plot size[b] [m^2]	Tenure length [years]	Construction Year	No. of bedrooms	Monthly Rent[a] [EUR]	Current Value[a] [EUR]
All tenure stati.							
Median	125.00	500.00	10.00	1985	3.00	1,500.00	600,000.00
Mean	139.48	1,028.99	15.32	1974	3.39	1,696.00	642,660.00
Std. Dev.	80.93	7,663.39	14.93	40.37	1.79	885.48	457,772.17
Owner-occupiers.							
Median	146.80	500.00	15.00	1984	4.00	1,800.00	652,000.00
Mean	163.54	1,104.31	18.91	1974	3.89	1,995.50	747,150.96
Std. Dev.	81.54	8,218.12	15.81	40.19	1.73	849.29	432,997.30
Renters.							
Median	80.00	345.20	4.00	1985	2.00	950.00	350,000.00
Mean	85.93	531.42	7.33	1974	2.28	1,030.56	410,048.00
Std. Dev.	46.60	551.83	8.34	40.80	1.37	530.78	424,992.38

Notes: The table reports summary statistics related to the household main residence (HMR). Measures are derived respecting survey weights.
[a] Prices and rents are reported by survey participants and hence are identified as ORS, RPR, RRS and ORR.
[b] The plot size is only available for houses and not for apartments.
Source: LU-HFCS, 3rd wave.

Table O.3 – Summary Statistics: Dwelling Characteristics – Advertisements

	Surface $[m^2]$	Plot size $[m^2]$	Construction Year	No. of bedrooms	Monthly Rent [EUR]	Sales Price [EUR]
All Dwelling Types.						
Median	97.00	0.00	1968	2.00		
Mean	118.50	8.05	1979	2.48		
Std. Dev.	74.58	485.07	30.02	1.36		
Dwellings for Sale.						
Median	113.58	0.00	1968	3.00		618,977.00
Mean	133.58	11.77	1979	2.81		702,564.00
Std. Dev.	77.19	597.22	32.05	1.35		397,664.35
Dwellings for Rent.						
Median	79.51	0.00	1968	2.00	1,500.00	
Mean	89.33	0.86	1979	1.85	1,684.00	
Std. Dev.	59.18	15.23	25.65	1.11	858.13	

Notes: The advertisements refer to the period 01 January to 31 December 2018. Overall there are 7,860 rent and 14,759 sales advertisements. For both categories, the overwhelming majority of properties do not come with any outside plots, which explains the median values for this variable.
Source: Housing Observatory.

Table O.4 – Distributional Wealth Measures – Totals.

	Reported [Million EUR]	Imputed [Million EUR]
Complete Sample ($N = 1,616$, $N_w = 226,378$)		
Net Wealth		
Total	203,273	210,060
Breakdowns by		
Net Income – Q1	17,805	20,253
Net Income – Q2	24,818	26,671
Net Income – Q3	28,436	30,345
Net Income – Q4	42,404	43,906
Net Income – Q5	89,811	88,885
Net Wealth – Q1	420	519
Net Wealth – Q2	7,917	11,599
Net Wealth – Q3	22,541	26,394
Net Wealth – Q4	38,666	40,674
Net Wealth – Q5	133,729	130,874
Owner-Occupiers ($N = 1,207$, $N_w = 156,210$)		
Net Wealth		
Total	190,504	197,290
Breakdowns by		
Net Income – Q1	16,985	19,434
Net Income – Q2	22,197	24,050
Net Income – Q3	26,658	28,566
Net Income – Q4	39,664	41,166
Net Income – Q5	85,000	84,075
Net Wealth – Q1	-56	44
Net Wealth – Q2	5,471	9,153
Net Wealth – Q3	21,051	24,904
Net Wealth – Q4	37,555	39,562
Net Wealth – Q5	126,483	123,627

Notes: Statistics supplementing Table 7.
Source: LU-HFCS, 3rd wave and authors' calculations based on data from the *Housing Observatory*, on advertisements available between 1 January and 31 December 2018.

Table O.5 – Adjusted Breakdowns for European Countries.

DE – Germany				
Weighted number of households:				17,733,246
Number of individual survey observations:				2,923
	Residential Housing Wealth		Net Wealth	
	Reported [mEUR]	Adjusted [mEUR]	Reported [mEUR]	Adjusted [mEUR]
Total	4,694,074	4,961,354	8,153,560	8,420,840
Breakdown by				
Gross Income – IQ1	227,017	232,521	333,343	338,848
Gross Income – IQ2	530,170	538,683	973,477	981,990
Gross Income – IQ3	780,210	791,999	1,085,329	1,097,118
Gross Income – IQ4	1,078,282	1,086,400	1,652,811	1,660,928
Gross Income – IQ5	2,078,395	2,065,919	4,108,600	4,096,123
Net Wealth – WQ1	36,039	36,076	-31,570	-31,533
Net Wealth – WQ2	35,781	36,916	6,181	7,316
Net Wealth – WQ3	377,320	389,698	273,465	285,844
Net Wealth – WQ4	1,261,479	1,283,173	1,488,129	1,509,822
Net Wealth – WQ5	2,983,454	2,906,437	6,417,355	6,340,337

FR – France				
Weighted number of households:				16,966,548
Number of individual survey observations:				9,613
	Residential Housing Wealth		Net Wealth	
	Reported [mEUR]	Adjusted [mEUR]	Reported [mEUR]	Adjusted [mEUR]
Total	3,955,517	4,180,744	6,459,271	6,684,498
Breakdown by				
Gross Income – IQ1	278,907	285,669	393,229	399,992
Gross Income – IQ2	402,967	409,438	519,985	526,456
Gross Income – IQ3	683,494	693,822	907,645	917,973
Gross Income – IQ4	949,263	956,409	1,342,903	1,350,049
Gross Income – IQ5	1,640,886	1,631,036	3,295,508	3,285,658
Net Wealth – WQ1	25,920	25,947	-4,250	-4,223
Net Wealth – WQ2	165,628	170,882	47,262	52,515
Net Wealth – WQ3	667,087	688,972	565,841	587,725
Net Wealth – WQ4	1,133,735	1,153,232	1,374,388	1,393,885
Net Wealth – WQ5	1,963,146	1,912,468	4,476,030	4,425,351

Table O.6 – The Source of deviations between reported and imputed values (cont'd).

IT – Italy				
Weighted number of households:				17,470,716
Number of individual survey observations:				5,337
	Residential Housing Wealth		Net Wealth	
	Reported [mEUR]	Adjusted [mEUR]	Reported [mEUR]	Adjusted [mEUR]
Total	3,570,757	3,774,076	5,221,735	5,425,053
Breakdown by				
Gross Income – IQ1	345,425	353,801	388,397	396,773
Gross Income – IQ2	447,415	454,600	511,246	518,431
Gross Income – IQ3	636,095	645,706	800,110	809,721
Gross Income – IQ4	825,960	832,178	1,198,805	1,205,023
Gross Income – IQ5	1,315,862	1,307,963	2,323,176	2,315,277
Net Wealth – WQ1	4,451	4,456	-21	-17
Net Wealth – WQ2	209,179	215,814	164,317	170,952
Net Wealth – WQ3	592,594	612,035	648,569	668,009
Net Wealth – WQ4	957,726	974,196	1,160,534	1,177,004
Net Wealth – WQ5	1,806,806	1,760,163	3,248,336	3,201,693

SK – Slovakia				
Weighted number of households:				1,644,998
Number of individual survey observations:				1,911
	Residential Housing Wealth		Net Wealth	
	Reported [mEUR]	Adjusted [mEUR]	Reported [mEUR]	Adjusted [mEUR]
Total	139,525	147,469	189,168	197,112
Breakdown by				
Gross Income – IQ1	15,915	16,301	17,720	18,106
Gross Income – IQ2	24,315	24,705	28,108	28,499
Gross Income – IQ3	24,717	25,090	30,228	30,601
Gross Income – IQ4	31,547	31,784	39,296	39,533
Gross Income – IQ5	43,031	42,773	73,816	73,557
Net Wealth – WQ1	6,712	6,719	2,615	2,622
Net Wealth – WQ2	15,949	16,455	13,968	14,474
Net Wealth – WQ3	23,954	24,740	25,426	26,212
Net Wealth – WQ4	33,036	33,604	39,398	39,967
Net Wealth – WQ5	59,873	58,328	107,760	106,214

Notes: Reported and adjusted residential housing and total net wealth breakdowns for owner-occupiers in European countries for groups formed by net wealth and gross household income quintiles. Adjusted figures are obtained by applying the measured shifts found for Luxembourg to breakdowns found for other countries.
Source: HFCS, 3rd wave.

Table O.7 – Demographic and Socio-Economic Structure

	All tenure stati		***Owner-occupier***		***Renter***	
Household						
	Mean	Median	Mean	Median	Mean	Median
Gross Income [EUR]	93,111	71,120	106,525	82,896	63,249	48,262
Net Wealth [EUR]	897,938	498,454	1,219,532	732,360	181,990	23,048
Size [No.]	2.4		2.5		2.1	
Dependant Children [No.]	0.6		0.7		0.6	
Reference Person						
	Share					
Employment status						
Employed	59%		54%		69%	
Self-employed	4%		4%		4%	
Unemployed	3%		1%		5%	
Retired	27%		34%		13%	
Other	7%		7%		9%	
Marital status						
Single	27%		21%		39%	
Couple	50%		56%		38%	
Divorced	14%		12%		20%	
Widowed	9%		11%		4%	
Education						
Low-education	26%		23%		30%	
Mid-education	38%		40%		35%	
High-education	36%		36%		35%	
Sex						
Female	42%		43%		40%	
Male	58%		57%		60%	

Notes: The table reports demographic and socio-economic characteristics for households in Luxembourg (residents). Measures are derived respecting survey weights. Individual characteristics are provided for the survey reference person, i.e., the main interview partner.

Source: LU-HFCS, 3rd wave.

Table O.8 – The Source of Deviation – Suppressed Output

Over-reporting

	(1)		(2)		(3)		(4)		(5)	
Over-reporting (reported > imputed) by	10%	15%	10%	15%	10%	15%	10%	15%	10%	15%
Sex										
Male					*(ref. cat.)*					
Female			1.028	1.078					1.053	1.100
			(0.179)	(0.189)					(0.191)	(0.199)
Age										
16-34					*(ref. cat.)*					
35-44			0.876	1.355					0.853	1.294
			(0.240)	(0.414)					(0.240)	(0.413)
45-54			0.896	1.222					0.806	1.033
			(0.252)	(0.367)					(0.239)	(0.329)
55-64			0.968	1.266					0.801	0.961
			(0.282)	(0.411)					(0.252)	(0.341)
65+			1.066	1.258					0.837	0.857
			(0.338)	(0.410)					(0.306)	(0.316)
Household size			1.064	1.025					1.054	0.995
			(0.072)	(0.071)					(0.077)	(0.075)
Country of birth										
Other					*(ref. cat.)*					
Luxembourg			1.325	1.518**					1.185	1.262
			(0.237)	(0.276)					(0.229)	(0.248)
Mode of acquisition (HMR)										
purchased or own construction					*(ref. cat.)*					
inherited or gifted					1.190	1.689			1.054	1.463
					(0.468)	(0.639)			(0.453)	(0.592)
Under-reporting (reported < imputed)										
Sex										
Male					*(ref. cat.)*					
Female			0.983	1.010					0.923	0.943
			(0.162)	(0.154)					(0.158)	(0.151)
Age										
16-34					*(ref. cat.)*					
35-44			1.042	1.279					1.056	1.269
			(0.251)	(0.282)					(0.268)	(0.295)
45-54			1.445	1.626**					1.476	1.539*
			(0.351)	(0.342)					(0.387)	(0.348)
55-64			1.365	1.539*					1.216	1.208
			(0.348)	(0.353)					(0.337)	(0.304)
65+			1.817**	1.806**					1.405	1.113
			(0.507)	(0.448)					(0.449)	(0.320)
Household size			1.011	0.997					1.000	0.972
			(0.062)	(0.055)					(0.068)	(0.061)
Country of birth										
Other					*(ref. cat.)*					
Luxembourg			0.769	0.744*					0.759	0.657**
			(0.127)	(0.118)					(0.140)	(0.122)
Mode of acquisition (HMR)										
purchased or own construction					*(ref. cat.)*					
inherited or gifted					1.148	1.632			1.224	1.816
					(0.421)	(0.600)			(0.482)	(0.721)

Notes: The table reports suppressed estimation results complementing Table 6.
Source: LU-HFCS, 3rd wave and authors’ calculations.

Table O.9 – Prices versus Rents versus Income

	Quantile Level – ϑ				
	10	25	50	75	90
All tenure stati.					
Price-to-Rent Ratio (PR)	24.7	27.6	32.0	37.1	42.6
Price-to-Income Ratio (PI)	6.4	8.4	12.0	18.7	31.3
Rent-to-Income Ratio (RI)	0.19	0.26	0.38	0.59	0.96
Owner-occupier.					
Price-to-Rent Ratio (PR)	24.9	28.0	32.8	37.8	43.9
Price-to-Income Ratio (PI)	6.3	8.1	11.4	17.5	30.6
Rent-to-Income Ratio (RI)	0.18	0.24	0.35	0.53	0.94
Renter.					
Price-to-Rent Ratio (PR)	24.3	26.7	30.5	34.9	39.3
Price-to-Income Ratio (PI)	6.8	9.4	13.5	21.3	32.0
Rent-to-Income Ratio (RI)	0.23	0.32	0.45	0.68	1.00

Notes: Numbers based on imputed objectified prices and rents. Ratios are computed according to formulae (PR), (PI) and (RI), respecting survey weights. The median is represented by $\vartheta = 0.5$.
Source: Authors' calculations using data from the *LU-HFCS*, 3rd wave and the *Luxembourg Observatoire de l'Habitat*.

BANQUE CENTRALE DU LUXEMBOURG
EUROSYSTÈME

2, boulevard Royal
L-2983 Luxembourg

Tél.: +352 4774-1
Fax: +352 4774 4910

www.bcl.lu • info@bcl.lu