TÜRKİYE CUMHURİYET MERKEZ BANKASI

Identification of Wealthy Households from the Residential Property Price Index Database for Sample Selection for Household Surveys

Evren CERİTOĞLU
Özlem SEVİNÇ

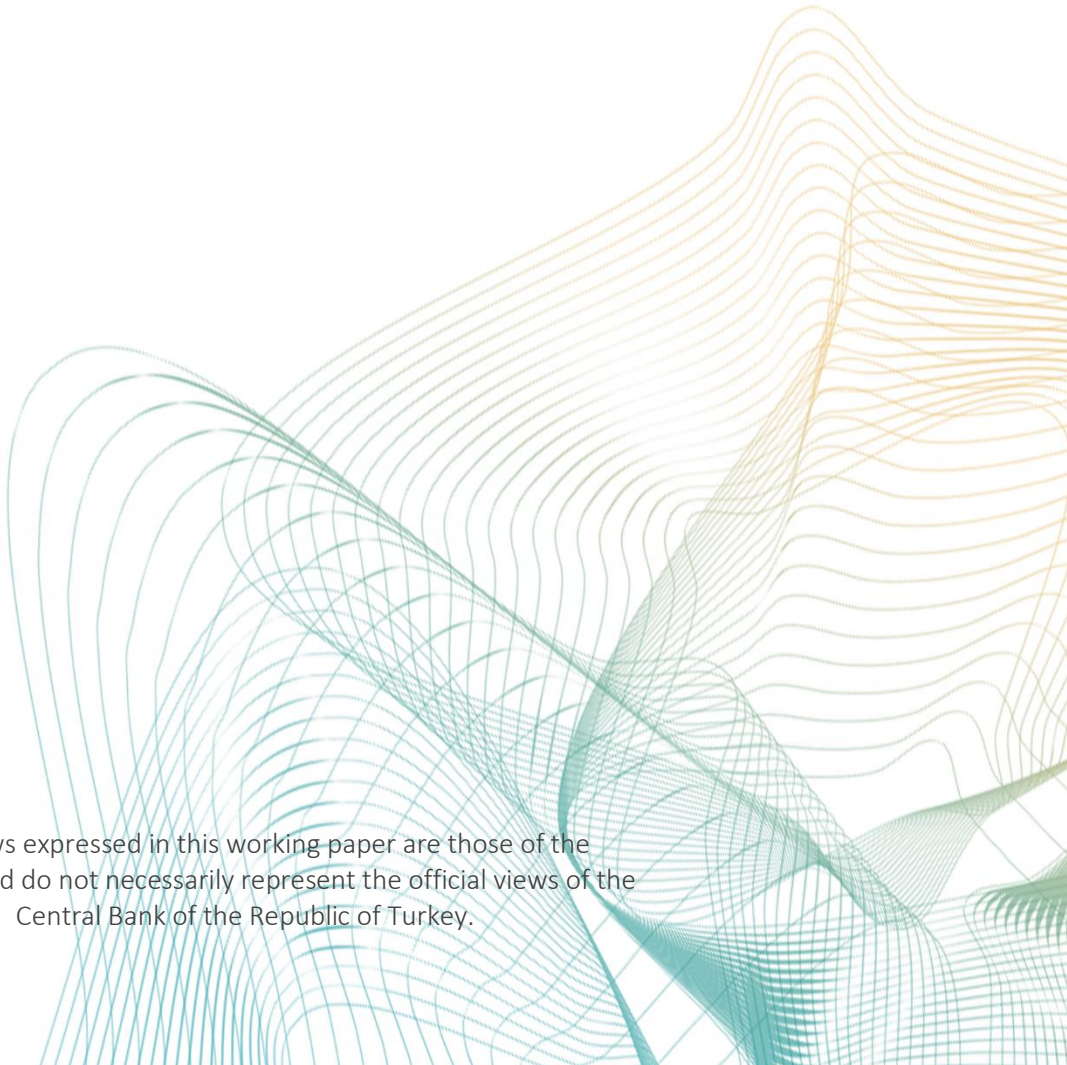
October 2020

Working Paper No: 20/10

Address:
Central Bank of the Republic of Turkey
Head Office
Structural Economic Research Department
Hacı Bayram Mh. İstiklal Caddesi No: 10
Ulus, 06050 Ankara, Turkey

Phone:
+90 312 507 80 04

Facsimile:
+90 312 507 78 96



Identification of Wealthy Households from the Residential Property Price Index Database for Sample Selection for Household Surveys [*]

Evren Ceritoğlu [a] and Özlem Sevinç [b]

Abstract

This paper aims to identify wealthy households in Turkey for sample selection for household surveys. In the absence of income and wealth tax data, we analyze house prices from the Residential Property Price Index (RPPI), which is constructed by the Central Bank of the Republic of Turkey (CBRT) from dwelling appraisal reports to monitor price movements. RPPI is announced monthly by the CBRT for Turkey and 26 geographical regions at NUTS2 level since 2012, but data actually starts from January 2010. The RPPI database comprises more appraisal observations from İstanbul and western provinces, where house prices are significantly higher than country average. However, the number of appraisal observations is low for the Eastern provinces, since the number of house sales is limited in poor and small provinces. Moreover, the percentage of mortgaged house sales is even lower in these regions, whereas the RPPI database is based on dwelling appraisal reports on house sales, which are subject to mortgage loans.

We examine unit house prices from the CBRT – RPPI database from 2010 to 2018 at province, district and neighborhood levels. Unit house prices are calculated by dividing the value (TL) to the gross usage area (m^2) at current prices. Only neighborhoods with 30 or more observations are examined in the analysis. We discuss the validity of the hypothesis that there is a direct relationship between unit house prices and the number of home appraisals. We regress the natural logarithm of the number of home appraisals on the natural logarithm of unit house prices using mean values. We perform fixed effects regressions at both neighborhood and province levels using our unbalanced and balanced panel data sets. We control for year effects by introducing time dummy variables into the regressions. We find that there is a positive and statistically significant relationship between unit house prices and the number of home appraisals. Moreover, we perform the same regressions for

[*] We would like to thank the editor, the anonymous referee, Özgül Atılgan Ayanoğlu, Duygu Konukçu Çelik, Ezgi Deryol and Erdi Kızılkaya for their contribution to this paper.

[a] Economist, Structural Economic Research Department, Central Bank of the Republic of Turkey (CBRT), Ulus / Ankara 06050 TURKEY
Telephone no: + 90 312 5078024
Fax number: + 90 312 5075732
E-mail: evren.ceritoglu@tcmb.gov.tr
[b] Assistant Specialist, Structural Economic Research Department, Central Bank of the Republic of Turkey (CBRT), Ulus / Ankara 06050 TURKEY
Telephone no: + 90 312 5078053
Fax number: + 90 312 5075732
E-mail: ozlem.sevinc@tcmb.gov.tr

neighborhoods that have more than 100 observations as a robustness check. We observe that the size and the sign of the regression coefficients do not change when we restrict our data set.

The direction of the relationship might be from the number of home appraisals to unit house prices or it could be both ways. For that reason, as another robustness check, we regress the natural logarithm of unit house prices on the natural logarithm of the number of home appraisals. We observe that there is a statistically significant relationship between the number of home appraisals and unit house prices. However, the size of the regression coefficients is considerably lower in this case. As a result, our empirical analysis indicates that the number of observations is higher in administrative units, where house prices are higher. Therefore, we argue that identification of wealthy households according to their neighborhoods using the RPPI database is a reliable and consistent method for oversampling for household surveys in Turkey.

***Key words**: Unit house prices, wealthy households, panel data, sampling design, oversampling*

***JEL codes**: C33, C83, R21, R31, R32*

Non–Technical Summary

This paper proposes a method to identify wealthy households for oversampling them on a neighborhood basis in household surveys in Turkey. In an ideal world, the results of a household survey represent the entire population. However, in practice, especially wealth-related surveys fail to interview with wealthy households. It is of great importance that they are represented in the sample in a balanced manner, since household assets and liabilities are mainly concentrated in the upper income groups. In such wealth-related surveys, an efficient application is to oversample wealthy households, which is about contacting proportionally more wealthy households in the surveys.

Income and wealth tax data at the individual or household level is the most appropriate source of information to identify wealthy households. However, income and wealth tax are not available at individual or household level in Turkey. Moreover, there is no direct data related to the wealthiest people's addresses including their provinces, districts and neighborhoods. Therefore, we analyze unit house prices for identifying wealthy households in Turkey, which is derived from the Central Bank of the Republic of Turkey (CBRT) – Residential Property Price Index (RPPI) database between 2010 and 2018 at province, district and neighborhood levels. One positive aspect of the RPPI database is that it is nationally representative, while it provides information about geographical regions at NUTS2 level. Another positive aspect of accessing house price information at the neighborhood level is obtaining a tool that can speak with the Turkish Institute of Statistics (TURKSTAT) sampling frame.

Housing wealth is often the largest component of household wealth. Moreover, household income and housing wealth are directly related to each other. Therefore, we assume that wealthy families live in more expensive neighborhoods in this paper. We argue that identification of wealthy households according to their neighborhoods using the RPPI database is a reliable and consistent method for oversampling them in household surveys in Turkey.

Unit house prices are calculated by dividing the value of the residence (TL) to the gross usage area (m^2) at current prices. We demonstrate that the distribution of neighborhoods and provinces with respect to unit house prices is very similar to income distribution across country. We also find that there is a positive and statistically significant relationship between unit house prices and the number of house sales at neighborhood and province levels. Thus, the empirical analysis confirms that our hypothesis is valid for the Turkish economy. The implementation of this method will be considered as an innovation, since TURKSTAT has not previously conducted a sampling design, which enables oversampling of wealthy households in their surveys.

I. Introduction

The aim of this paper is to develop a reliable method to identify wealthy households for sampling design for household surveys in Turkey. Wealthy households hold a larger share of financial assets and liabilities. Moreover, they own a higher variety of financial assets and liabilities (Causa *et al.*, 2019). Similarly, Bertaut and Starr-McCluer (2002) states that ownership of variates of financial assets and liabilities increases with wealth except for credit card balances and some kinds of debt in the U.S. economy. They also argue that there is large gap between intensity of assets and liabilities over different wealth groups. While aiming a survey on assets and liabilities, it is necessary to approach as rich households as possible to accurately represent the complete distribution of wealth (Balestra and Tonkin 2018; Vermeulen, 2016 and 2018). In this context, it would be better to use information from administrative data to oversample households that are wealthy.

Previous empirical literature suggests that income and wealth tax data at the individual or household level is the most appropriate source of information to identify wealthy households (Bricker *et al.*, 2016). However, income and wealth tax are not available at the individual or household level in Turkey. For that reason, the sampling strategy must focus on finding a variable that will reflect household wealth in the most detailed level according to available resources. Moreover, it should be possible to match the selected proxy variable with the sampling frame of the Turkish Institute of Statistics (TURKSTAT). As a result, in the absence of tax data, we analyze house prices from the Residential Property Price Index (RPPI), which is constructed by the Central Bank of the Republic of Turkey (CBRT) from dwelling appraisal reports to monitor price movements in Turkey.[1] One of the positive aspects of the RPPI database is that it is nationally representative and it provides information about geographical regions at NUTS2 level. Another positive aspect of accessing house price information at the neighborhood level is obtaining a tool that can speak with TURKSTAT's sampling frame.[2]

Housing wealth is often the largest component of household wealth. Moreover, household income and housing wealth are directly related to each other. For that reason, we assume that wealthy families live in more expensive neighborhoods. In particular, we discuss the validity of the hypothesis that there is a direct relationship between house prices and the number of house sales. We perform econometric tests using unit house prices and the number of appraisal reports from the RPPI database at both neighborhood and province levels using balanced and unbalanced panel data sets to test this

[1] https://www.tcmb.gov.tr/wps/wcm/connect/EN/TCMB+EN/Main+Menu/Statistics/Real+Sector+Statistics/Residential+Property+Price+Index/

[2] TURKSTAT is one of the exceptional institutions that has access to the addresses of households in Turkey and has the authority to provide these information for household surveys carried out by institutions other than them with official requests under certain conditions.

hypothesis. The establishment of such a relationship will indicate that the RPPI database is sufficient to identify wealthy households for sample selection for household surveys in Turkey.

The main contribution of this paper is to show that unit house prices successfully predict the spatial distribution of income in Turkey and can be used for sample selection on a neighborhood basis. Accordingly, first we demonstrate that the distribution of neighborhoods and provinces with respect to unit house prices is very similar to income distribution across country, which is measured using both aggregate and micro-economic data. Second, we find that there is a positive and significant association between the number of home appraisals and unit house prices at both neighborhood and province levels by performing econometric estimations using balanced and unbalanced panel data sets. These empirical findings suggest that the number of home transactions are higher in wealthy neighborhoods. Thus, we can argue that the small number of observations in the RPPI database from poor regions is not a major obstacle in the mapping of wealthy neighborhoods. Third, we control for the roles of income per capita, housing supply and population growth in the relationship between the number of home appraisals and unit house prices at province level in both unbalanced and balanced panel data estimations as a robustness check. We confirm that the positive relationship between the number of home appraisals and unit house prices is robust to the inclusion of control variables in the empirical analysis. As a result, we conclude that unit house prices can also be used for oversampling of wealthy households according to the proposed method in this paper, considering the importance of real estate ownership in the distribution of household wealth. Finally, the implementation of this method will be considered as an innovation, since TURKSTAT has not previously conducted a sampling design, which enables oversampling of wealthy households in their surveys.

The outline of the paper is as follows: Section II discusses the literature on sampling design in household surveys with a special emphasis on oversampling of wealthy households. Section III presents the theoretical background. Section IV provides a descriptive analysis of the RPPI database and section V presents the econometric results. Finally, section VI concludes this paper with a brief summary of our findings.

II. Sampling Design

In most of the countries coordinated by European Statistical Office (Eurostat), multistage stratified cluster sampling is adopted as the main sampling method for household surveys. In Turkey, TURKSTAT is also implementing this method successfully by using two stage stratified cluster sampling. In order to understand the sample selection, the method can be summarized as follows: A sampling frame, which includes information about all household addresses should be set and the coverage of

the population should be determined at the beginning. All households that live in Turkey are included in the TURKSTAT National Address Database (NAD) which is based on the Address Based Population Registry System, while the institutionalized population – individuals that live in dormitories, guesthouses, childcare centers, nursing homes, private hospitals, prisons and military barracks – is excluded from the sampling frame.[3] NAD is updated every six months to take into account the situation of moving people from one address to another address. The sites like villages including less than 20 households which are at most 1% of the country population are not covered in the sampling frame. The sample design strata are described by geographic regions (NUTS) and area types (urban and rural). The type of used geographic region can be changed from one survey to another. For example, if the implicit strata are defined as NUTS3 (including 81 provinces) and urban-rural areas, there exists 162 strata. The urban and rural definition comes from the number of the residents that live in a site. In the first stage, the primary sampling units (PSU) consisting approximately 100 addresses called blocks are derived from sampling frame. While PSU's are formed in the sites having municipality, villages become PSU's by themselves. The primary sampling units are selected by probability proportional to size (PPS) method by systematic sampling of PSU's ordered by geographical level (NUTS1, NUTS2, Province, District, etc.).[4] Then, final sampling units that are called households are selected systematically from primary sampling units.[5] The block system that TURKSTAT applies is very good compared with other countries, since the blocks are very small, which is a necessary and useful property that raises the efficiency of the probability sample selection stage. All major TURKSTAT household surveys including Household Budget Survey (HBS), Household Labor Force Survey (LFS) and Survey on Income and Living Conditions (SILC) have the same sampling methodologies as mentioned.

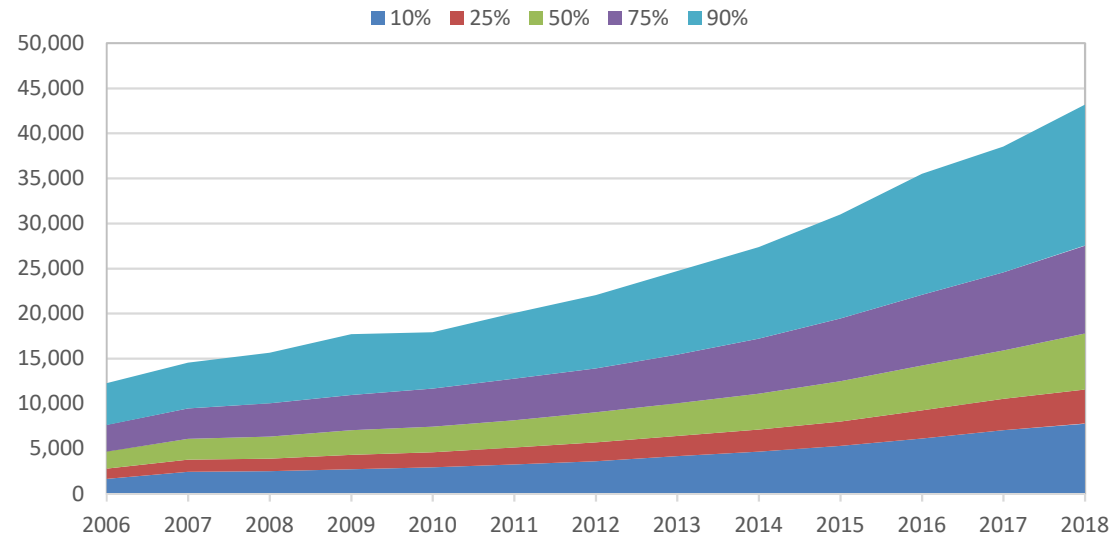
In an ideal world, the results of a survey study represent the entire population. On the other hand, in practice, especially wealth related surveys fail to interview with the wealthy households. Wealthy households are less willing to participate in the surveys. It is of great importance that they are represented in the sample in a balanced manner and that they give correct answers to the questions asked, since household assets and liabilities are mainly concentrated in the upper income group (Causa *et al.*, 2019).

We observe that the dispersion of household disposable income is significantly larger in upper income groups compared to lower income groups (Figure 1). The high degree of dispersion makes it difficult to estimate mean and median levels correctly in upper income groups. A randomly selected observation is closer to mean and median values in lower income groups, whereas a randomly selected observation could be significantly different from mean and median values in upper income groups. In

[3] http://www.turkstat.gov.tr/PreTablo.do?alt_id=1059
[4] Please see Appendix 1 for more information on geographical distribution.
[5] http://www.turkstat.gov.tr/UstMenu.do?metod=metabilgi

addition to that, non-response rate is generally higher among upper income groups compared to lower income groups in household surveys. As a result, at the sampling design stage, we need to select more observations from upper income groups to reach unbiased estimates not only for these groups, but also for whole population.[6]

Figure 1 – The Highest Annual Equivalised Household Disposable Income for Cumulative Percentage Income Groups * (Current Prices, TL)



Source: TURKSTAT SILC

* Household disposable income is adjusted for family size and intra-household resource allocation between adults and children.

In wealth related surveys, an efficient application is to oversample wealthy households, which is about contacting proportionally more wealthy households in the surveys (Chakraborty and Waltl, 2018). Moreover, in order for the sample to represent a consistent distribution of wealth in the population, it is important to have a higher proportion of wealthy households in the sample than the normal sample distribution (Kennickell, 2008). This approach will lead to more observations in a certain part of the distribution than calculated from the original sampling frame. The oversampling method can also be used for finding rare sub-populations (Kalton, 2009) and finding hard to reach segments of population such as homeless persons, drug users, victims of female circumcisions (Marpsat and Razafindratsima, 2010) beside reaching wealthy households.

[6] It is necessary to consider family size and intra-household resource allocation in calculating income distribution indicators such as Gini coefficient and poverty line. TURKSTAT uses OECD equivalence measure in all household surveys. OECD equivalence assumes the value of 1 for the reference person in the household, 0.5 for household members, who are 14 and older, and 0.3 for household members, who are younger than 14. Household disposable income is divided by OECD equivalence scale. Thus, it becomes possible to compare households with different sizes and types with each other.

In this case, the most accurate way seems to include the wealthiest people more than usual in the sample. Valliant *et al.* (2014) studied the use of artificial variables based on commercial sources for sampling as stratification to reach sub-groups in the population. There are many other ways in order to find the wealthiest individuals. Wealth and income tax data are just a few of these ways, which are most reliable and appropriate sources.

In Europe, HFCS (Household Finance and Consumption Survey) coordinated by the European Central Bank (ECB) has been practiced for a while, where the wealthy people are included more with the help of oversampling method.[7] The hardest part of finding the wealthy with the oversampling method is the necessity of having current information representing the entire population in the sampling frame. The more the variable used for the oversampling method and the stronger the relationship between the variables and the wealth, the more successful the results of the method at the end of the application. Countries are trying to apply the oversampling method within the framework of their data, which helps them to find wealthy people. Spain and France use personal taxable wealth data, which is the best indicator of wealth while Estonia, Latvia, Luxembourg and Finland use personal income as an indicator of wealth. Housing price is another indicator of wealth, which used by Belgium, Germany and Greece while Poland and Portugal use the property size for identifying the wealth. Household Finance and Consumption Network (HFCN, 2016a and 2016b) report that electricity consumption, regional income, personal education and labor status are some of the ways of finding wealthy households.

In Turkey, there is no direct data related to wealthiest people's addresses including their provinces, districts and neighborhoods, which are existing administrative units. On the other hand, it is stated that personal wealth tax data is the best indicator of wealth (Bricker *et al.*, 2016). Even if any variable exists that determines the wealth level of individuals in Turkey, there is no sampling frame to match this variable by individuals. The TURKSTAT sampling frame is based on household addresses rather than individuals. For this reason, the necessity of finding a variable, which indicates the wealthy, has arisen over the administrative units such as provinces, districts and neighborhoods in order to match with the sampling frame provided by TURKSTAT.

Administrative data that indicate wealthy households through administrative units have been examined in detail, taking into account similar country examples. It has been decided that unit house price is the most suitable indicator for identifying wealthy households, which is included in the CBRT – RPPI database from 2010 to 2018 at province, district and neighborhood level. Housing wealth is often

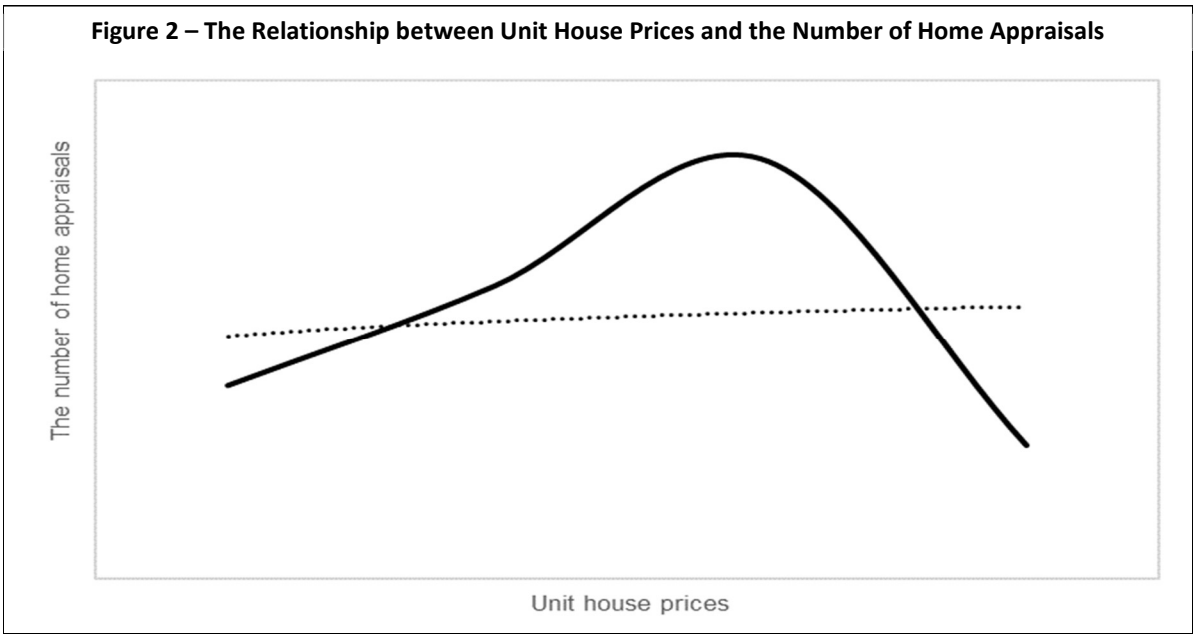
[7] https://www.ecb.europa.eu/pub/economic-research/research-networks/html/researcher_hfcn.en.html

the largest component of household wealth. For this reason, we assume that wealthy families live in more expensive neighborhoods in this paper.

III. Theoretical Background

From a theoretical point of view, we expect to find a positive relationship between unit house prices and the number of home appraisals under the assumption that housing supply is constant (Knoll *et al.*, 2017). However, we expect that the number of home appraisals will increase as unit house prices increase up to a certain point, but the number of home appraisals will begin to fall when unit house prices exceed a critical point. We think that there will be fewer observations for the wealthiest households for several reasons. The number of houses for sale may be lower in the most expensive neighborhoods. Moreover, the wealthiest households are less likely to apply for a housing loan for a home purchase. Thus, we predict that this relationship will have a concave shape (Figure 2).

Figure 2 – The Relationship between Unit House Prices and the Number of Home Appraisals



House prices and house sales are expected to be higher in residential areas where housing demand is strong (Rosen and Smith, 1983; Riddel, 2004; Steiner, 2010). Previous empirical literature shows that there is a positive and significant relationship between housing demand and household permanent income (Goodman, 1998 and 1990; Zabel, 2004). Moreover, the housing market is an aggregation of many local housing markets, which necessitates that the empirical analysis is carried out for smaller administrative units (Kiel and Zabel, 2008).

Similarly, previous studies on the Turkish economy find a positive and significant relationship between housing demand and household permanent income (Halicioglu, 2007; Ceritoğlu, 2017 and 2020). In particular, Ceritoğlu (2017) finds that house price changes have a positive and significant effect on the growth of cohort consumption in Turkey. He constructs a pseudo-panel data set using birth-year cohorts from twelve consecutive waves of HBS between 2003 and 2014. According to his findings, homeowners perceive their housing wealth higher as house prices rise, which affects their consumption decisions positively. Thus, his empirical findings support the wealth channel argument in explaining the relationship between house prices and household consumption. Moreover, Ceritoğlu (2020) estimates that the permanent income elasticity of housing demand is approximately 26% analyzing fourteen consecutive waves of HBS from 2003 to 2016. In the case of Turkey, there was excess supply in the housing market across country and in three major provinces throughout the period of analysis.[8]

In the equations unit house prices and the number of home appraisals are denoted by *U* and *A*, respectively. Neighborhood or province is shown by *i*, while year is shown by *t* in the equations from (1) to (4).

$$U_{it} = \alpha_{it} + \beta_1 A_{it} + \nu_{it} \quad (1)$$

$$A_{it} = \gamma_{it} + \delta_1 U_{it} + \upsilon_{it} \quad (2)$$

Moreover, *Z* represents social and demographic variables, which are used as control variables in the equations (3) and (4). Finally, *v* and *υ* denote error terms in the equations from (1) to (4).

$$U_{it} = \alpha_{it} + \beta_1 A_{it} + \beta_2 Z_{it} + \nu_{it} \quad (3)$$

$$A_{it} = \gamma_{it} + \delta_1 U_{it} + \delta_2 Z_{it} + \upsilon_{it} \quad (4)$$

The selected control variables in this context are population, income, building permits and the number of households in a province. The estimations will be carried out for balanced and unbalanced panel data sets at the neighborhood and province levels in the empirical analysis section. They will also include time dummy variables to control for macro-economic effects that might have taken place in the period of analysis.

[8] https://tcmbblog.org/wps/wcm/connect/blog/en/main%20menu/analyses/what%20does%20the%20housing%20supply%20tell%20us

IV. Data

IV.1. CBRT – RPPI Database

The CBRT publishes RPPI for Turkey and 26 geographical regions at NUTS2 level on a monthly basis since 2012, but data actually starts from January 2010. RPPI is the only nationally representative house price index for the Turkish economy. Moreover, RPPI is produced as a hedonic house price index, which is adjusted for quality growth using construction properties (Hülagü *et al.*, 2016).

RPPI is based on dwelling appraisal reports, which are prepared by private appraisal firms at the request of deposit banks if a loan is demanded for a house sale. An appraisal report is prepared by professional appraisers for all houses whether a sale with bank loan takes place or not in the end. Appraisal reports are considered as a more reliable source of information about house prices, since asking prices that are put forward by sellers might be higher than actual market values. Moreover, households might under-report transaction prices at the Land Registry Offices to avoid real estate tax. In addition to that, RPPI database is not restricted to houses that are sold, which helps to avoid a potential sample selection bias.

A neighborhood is an administrative unit, which is a sub-region of a district and a district is a sub-region of a province in Turkey.[9] The primary sampling units consisting approximately 100 addresses are called blocks and are derived from sampling frame as TURKSTAT classification. However, a neighborhood, which is significantly larger than a block, consists of 700 household addresses on average. The fact that the blocks are smaller than the neighborhoods is a factor that increases the power of probability sampling.

In certain cases neighborhood names are either written wrongly in the appraisal reports or they are old names of these neighborhood, which are not used in public records anymore. At the same time new neighborhoods are constantly formed, since urban regions are growing swiftly due to population growth. As a result, we cross-checked neighborhood names from the RPPI database with neighborhood codes from the NAD, which is used as *sampling frame* by the TURKSTAT during the sampling design stage of household surveys, prior to our empirical analysis. However, this process led to the decline of the number of available neighborhood observations from the RPPI database. Despite the fall in the number of neighborhood observations, we observe that the RPPI database has a high degree of representation of household addresses for the country as a whole. Moreover, representation capacity of the RPPI database is especially larger in the Western regions, where house prices and house sales are significantly higher than the rest of the country (Table 1).

[9] For instance, İstanbul is a province in Turkey. İstanbul is also a separate region both in NUTS1 and NUTS2 classifications due to its large population and economic size. Ümraniye is a district of İstanbul, whereas Esenkent is a neighborhood of Ümraniye (see Appendix 2).

In this paper, we analyze unit house prices and the number of home appraisals from the RPPI database at the neighborhood level from 2010 to 2018. However, unit house prices and the number of home appraisals are available only annually at the neighborhood level. Unit house prices are calculated by dividing the value of the residence (TL) to the gross usage area (m^2) at current prices. Moreover, only neighborhoods with 30 or more observations are included in our data set. We calculate unit house prices and the number of home appraisals for provinces from neighborhood observations by using the number of home appraisals as weights.

Table 1 – Regional Representation (*Block Numbers*, *2016-2018 Average*)

NUTS1	National Address Database	Residential Property Price Index Database	(%)
TR1	43,478	38,742	89.11
TR5	24,319	18,883	77.65
TR2	10,966	7,996	72.92
TR4	23,409	17,025	72.73
TR3	33,813	21,910	64.80
TR8	11,624	6,671	57.39
TR6	29,831	17,119	57.39
TR7	10,381	5,808	55.95
TR9	7,615	3,462	45.46
TRA	3,988	1,788	44.83
TRC	17,830	7,557	42.38
TRB	7,959	3,152	39.60
Total	225,213	150,113	66.65

Source: TURKSTAT, CBRT and Authors' calculations

Our preliminary analysis shows that only İstanbul (TR1) and Aegean (TR3) regions have higher unit house prices than country average at NUTS1 level in 2018. In a similar fashion, İstanbul (TR10), İzmir (TR31) and Aydın-Denizli-Muğla (TR32) regions have higher unit house prices than country average at NUTS2 level in 2018 (Figure 3). We also observe that the ranking of geographical regions in terms of unit house prices do not change much between 2010 and 2018. In a similar fashion, when we analyze unit house prices with respect to districts, we observe that the majority of districts with the highest unit house prices are from İstanbul in successive years (Table 2). Highest unit house prices are measured consistently in Beşiktaş district of İstanbul over the years. The only exceptions are beautiful and popular seaside towns in İzmir, Muğla, and Antalya in some years.

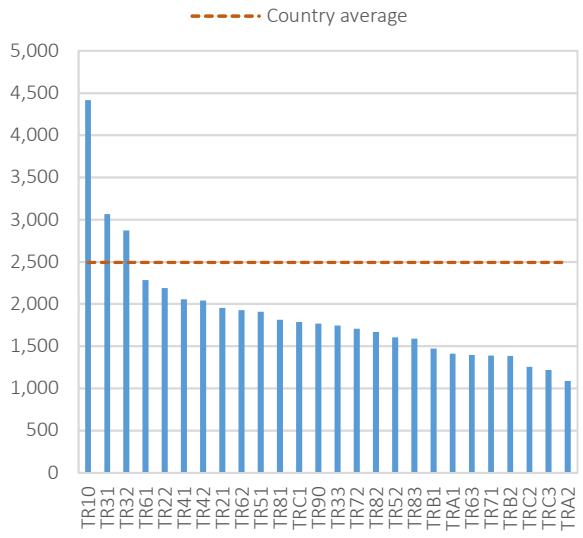
Table 2 – Unit House Prices with respect to Districts

	2016	2017	2018
1.	İstanbul Beşiktaş	İstanbul Beşiktaş	İstanbul Beşiktaş
2.	İstanbul Beykoz	İstanbul Beykoz	İstanbul Beykoz
3.	İstanbul Sarıyer	İstanbul Sarıyer	İstanbul Sarıyer
4.	İstanbul Kadıköy	İstanbul Bakırköy	İstanbul Bakırköy
5.	İstanbul Bakırköy	İstanbul Kadıköy	İstanbul Kadıköy
6.	İstanbul Şişli	İstanbul Şişli	İstanbul Şişli
7.	İstanbul Üsküdar	İstanbul Üsküdar	Muğla Bodrum
8.	İstanbul Ataşehir	Muğla Bodrum	İstanbul Üsküdar
9.	Muğla Bodrum	İstanbul Ataşehir	İzmir Çeşme
10.	İstanbul Zeytinburnu	İzmir Çeşme	İstanbul Ataşehir
11.	İzmir Çeşme	İstanbul Zeytinburnu	İstanbul Zeytinburnu
12.	İstanbul Maltepe	İstanbul Maltepe	İzmir Narlıdere
13.	İstanbul Ümraniye	İstanbul Ümraniye	İstanbul Kartal
14.	İstanbul Kartal	İstanbul Eyüp	İstanbul Maltepe
15.	İstanbul Fatih	İstanbul Kağıthane	İstanbul Ümraniye
16.	İstanbul Bahçelievler	İzmir Narlıdere	İstanbul Kağıthane
17.	İstanbul Kağıthane	İstanbul Kartal	İzmir Güzelbahçe
18.	Antalya Kaş	İstanbul Fatih	İstanbul Eyüp
19.	İstanbul Eyüp	İstanbul Bahçelievler	İstanbul Fatih
20.	İstanbul Bayrampaşa	İstanbul Adalar	İstanbul Bayrampaşa

Source: CBRT and Authors' calculations

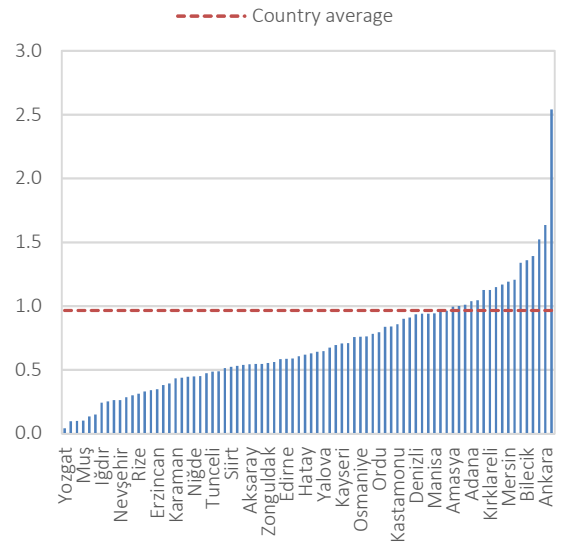
The ratio of the number of home appraisals to the number of household addresses, which was 1.78% in 2016 and 1.80% in 2017, decreased to 0.96% in 2018 across country (Figure 4). This ratio was higher than country average in 17 provinces, which were Amasya, Bartın, Bursa, Adana, Bolu, Uşak, Kırklareli, Kocaeli, **İstanbul**, Mersin, Gaziantep, Antalya, Bilecik, Eskişehir, Aydın, **Ankara** and Tekirdağ, in 2018. **İzmir** is included in this list in both 2016 and 2017. The Gross Domestic Product (GDP) per capita levels are also higher than the country average in these provinces in the period of analysis, which suggests that there is a positive relationship between house sales and income levels across provinces as expected. The fall in the number of home appraisals relative to the number of household addresses in 2018 stemmed from the contraction in the housing market. As the demand for housing credit declined due to the increase in interest rates, the number of home appraisal reports commissioned by deposit banks also fell in 2018. The ratios of the number of home appraisals to the number of household addresses indicate that every year at least 1% of all houses are appraised by private appraisal firms across country.

Figure 3 – Unit House Prices (TL / m^2) **across NUTS2 Regions** (2018)



Source: CBRT, Authors' calculations

Figure 4 – The Number of Home Appraisals and Household Addresses (2018)



Source: CBRT, Authors' calculations

We have more observations for the Western regions of the country, where both house prices and income levels are higher, from the RPPI database. However, we have few observations for small provinces, mainly in the Eastern parts of the country, where house prices and income levels are lower, since the number of house sales is limited (Table 1). Another important reason for the low number of observations in the Eastern provinces is that the percentage of mortgaged house sales is even lower in these regions, while the RPPI database is based on dwelling appraisal reports on house sales, which are subject to mortgage loans. Specifically, housing demand is stronger and house prices are higher in coastal cities. Approximately 32% of all house sales took place in three major provinces – İstanbul, Ankara and İzmir – in 2018, which explains the high number of observations from Western regions (Table A3.1). In contrast, in that year only 11% of all house sales took place in North East Anatolia (TRA), Middle East Anatolia (TRB) and South East Anatolia (TRC) regions, which are composed of 24 provinces.[10] As a result, the number of home appraisals are not evenly distributed throughout the country.

IV.2. The Relationship between Unit House Prices and the Number of Home Appraisals

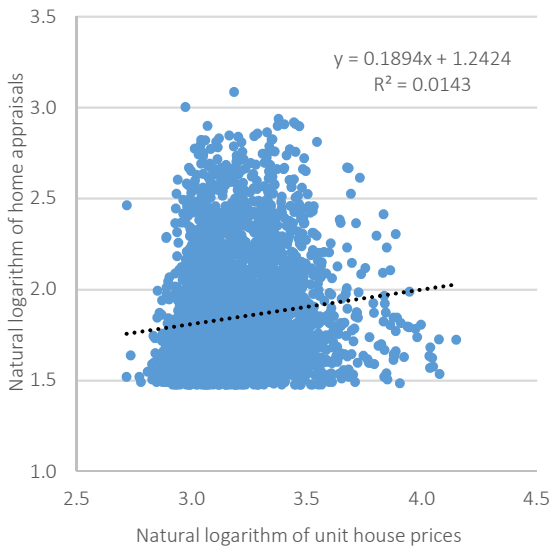
Our preliminary empirical analysis suggests that there is a positive relationship between unit house prices and the number of home appraisals. However, this analysis does not provide information

[10] The highest number of house sales were realized in İstanbul with 234,055 units in 2018. In comparison to that the lowest number of houses took place in Hakkari, Ardahan and Bayburt provinces with only 159, 189 and 543 units in 2018, respectively.

about the direction or the strength of the relationship between unit house prices and the number of home appraisals.

First, we take the natural logarithms of unit house prices and the number of home appraisals using 2010-2018 mean values. Other things being equal, we observe a positive, but weak relationship between unit house prices and the number of home appraisals at the neighborhood level (Figure 5). However, we observe a positive and stronger relationship between unit house prices and the number of home appraisals at the province level, other things being equal (Figure 6).

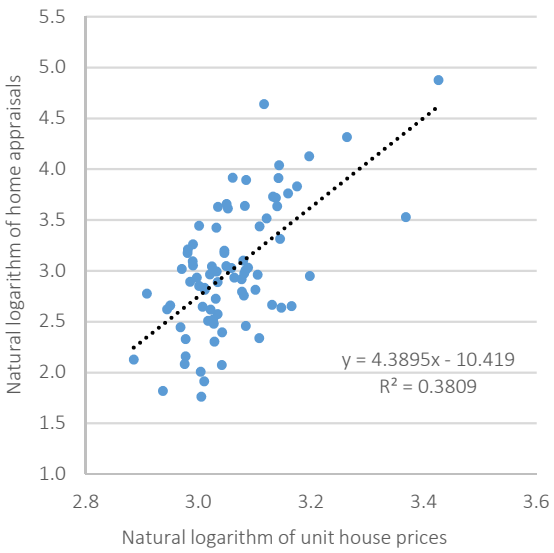
Figure 5 – Neighborhood (2010 – 2018 average) *



Source: CBRT, Authors' calculations

* Calculations are made for 2010-2018 period at neighborhood level using 3,930 observations.

Figure 6 – Province (2010 – 2018 weighted average) *

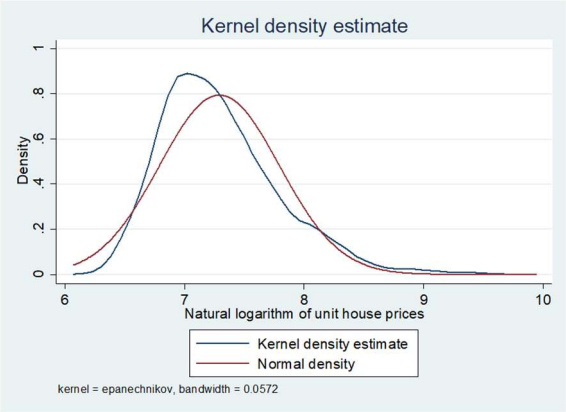


Source: CBRT, Authors' calculations

* Calculations are made for 2010-2018 period at province level using 81 observations.

We observe that the distribution of the number of home appraisals is more skewed to the right compared to unit house prices (Figure 7 and Figure 8). We also observe an upward time trend in both unit house prices and the number of home appraisals in the period of analysis across country (Figure 9). For that reason, we take the first differences of the natural logarithms of unit house prices and the number of home appraisals with respect to the previous year in the econometric estimations in the next section. Finally, when we analyze unit house prices and the number of home appraisals for the whole country, the decline in the number of home appraisal reports is more pronounced.

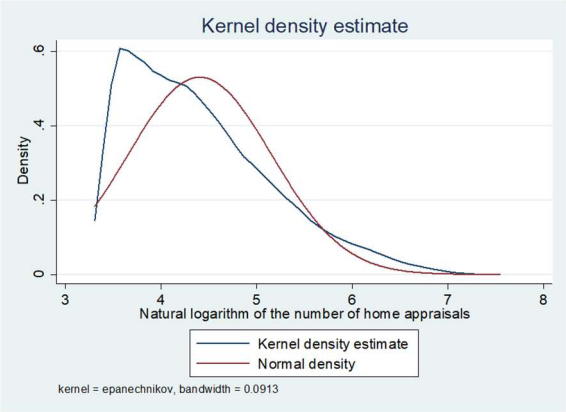
Figure 7 – Unit House Prices *



Source: CBRT, Authors' calculations

* Calculations are made for 2010-2018 period at neighborhood level using 3,930 observations.

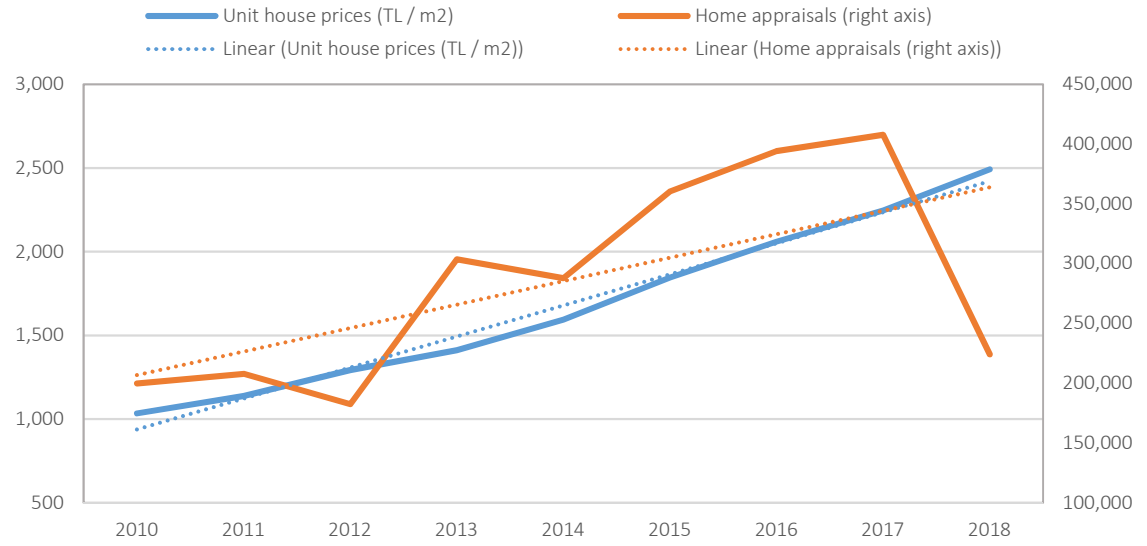
Figure 8 – The Number of Home Appraisals *



Source: CBRT, Authors' calculations

* Calculations are made for 2010-2018 period at neighborhood level using 3,930 observations.

Figure 9 – Turkey (2010 – 2018 weighted average)



Source: CBRT, Authors' calculations

IV.3. Oversampling

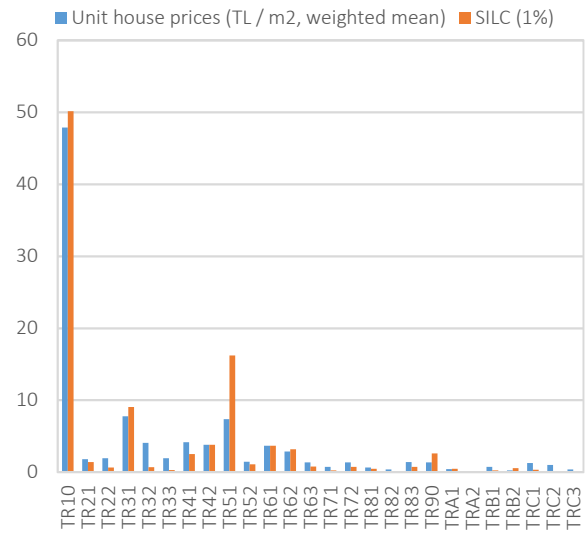
As it is stated previously, oversampling is basically used to increase the precision at the top of the income distribution by including more wealthy households than usual (Chakraborty *et al.*, 2019). At this stage, an extensive search for a suitable data source is carried out and it is decided to use unit house prices from the CBRT – RPPI database for oversampling in order to include more wealthy people than original sample.

It is realized that inconsistencies might occur for some regions only when unit house prices are used, so unit house prices are weighted with related geographical variable. The reason for the

weighting of the geographical regions with their respective shares in total population is to reflect their population density to unit house prices. For example, when the RPPI database is analyzed, it is observed that unit house prices in holiday regions such as Bodrum (Muğla) and Çeşme (İzmir) are very high, which may cause bias. At the same time the districts with the highest unit prices are mostly from İstanbul (Table 2). By using weighted RPPI ratio, we minimize the possibility that such observations affect the average unit house prices artificially.

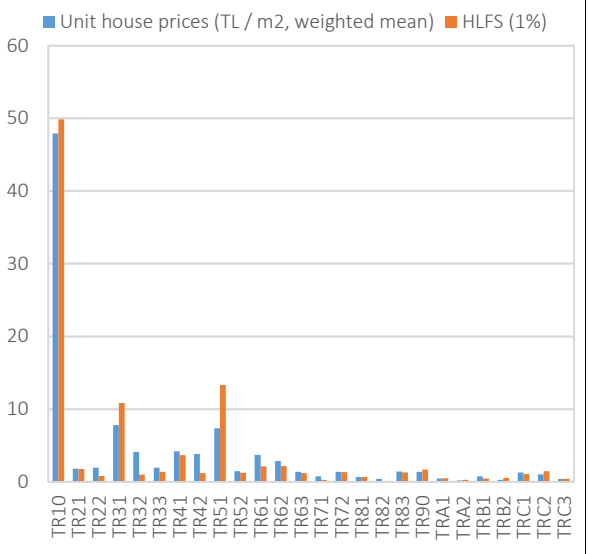
Beside discussing validity of the hypothesis that there is a direct relationship between unit house prices and the number of home appraisals, the suitability of the CBRT – RPPI database is also checked by different channels. Geographically (NUTS 2) weighted house prices are compared with the distribution of the top income group (%1) obtained from SILC and LFS at NUTS 2 level (Figure 10 and Figure 11). Especially for İstanbul, where the wealthiest segment of population of the country is located, reaching very close rates with Istanbul's share in the wealthiest income group from different household surveys, which are prepared by the TURKSTAT, indicates that it is appropriate to use unit house prices as an indicator of wealth in Turkey (Figure 10 and Figure 11).

Figure 10 – Survey on Income and Living Conditions (2016)



Source: TURKSTAT, CBRT, Authors' calculations

Figure 11 – Household Labor Force Survey (2017)



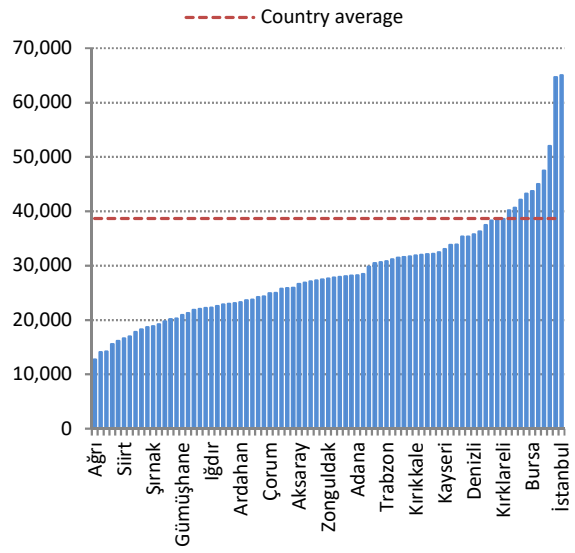
Source: TURKSTAT, CBRT, Authors' calculations

Our preliminary analysis using data from different sources give consistent results. İstanbul (TR10) comprises almost half of all highest income earners in Turkey. Moreover, Ankara (TR51) and İzmir (TR31) follow İstanbul in terms of income shares. After Adana-Mersin (TR61) and Antalya-Isparta-Burdur (TR62), the corresponding ratios are very low in the rest of the geographical regions (Central

Anatolia, Western Black Sea, Eastern Black Sea, Northeast Anatolia, Middle East Anatolia, Southeastern Anatolia), which suggests that the ratio of wealthy households in these regions is very low.

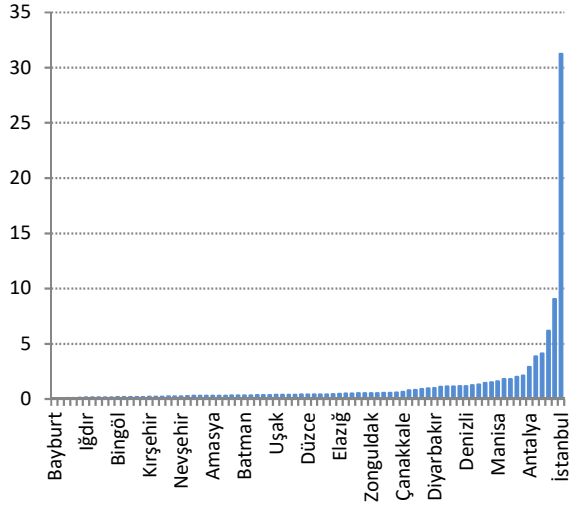
Regional unit house prices are calculated using 2016-2018 period averages for NUTS2 regions and weighted with their respective regional shares in total population. The value corresponding to İstanbul is very close to 50% in both regional unit house prices and the highest income quintile from 2016 wave of the Survey of Income Distribution and Living Conditions (SILC).[11] Moreover, the correlation coefficient between these two indicators is 0.98 (Figure 10). In a similar fashion, the highest income quintile from 2017 wave of the Household Labor Force Survey (LFS) is approximately 50% and the correlation coefficient between this indicator and regional unit house prices is 0.99 (Figure 11).[12]

Figure 12 – GDP per capita (Current prices, TL, 2017)



Source: TURKSTAT, CBRT, Authors' calculations

Figure 13 – Shares of Provinces in GDP (%, 2017)



Source: TURKSTAT, CBRT, Authors' calculations

Moreover, there are only 10 provinces, which have GDP per capita figures above country average in 2017, which are Bolu, Eskişehir, Yalova, Bilecik, Bursa, **İzmir**, Tekirdağ, **Ankara**, Kocaeli and **İstanbul**, respectively (Figure 12). We observe that the shares of provinces in GDP is even more unequally distributed. İstanbul accounts for 31.2% of GDP on her own, while Ankara and İzmir receive 9% and 6.2% of GDP in Turkey in 2017 (Figure 13).

[11] The distribution of the top 1% individual disposable income across NUTS2 regions is calculated using individual weights. 2016 SILC data refers to 2015 income.

[12] HLFS only reports labor income. The distribution of the top 1% individual labor income across NUTS2 regions is calculated using individual weights.

V. Econometric Results

V.1. Neighborhoods

We regress the first difference of the natural logarithm of the number of home appraisals on the first difference of the natural logarithm of unit house prices. We perform fixed effects regressions using our unbalanced panel data set. We also control for year effects in the regressions by introducing time dummy variables. We observe that there is a statistically significant relationship between unit house prices and the number of home appraisals at the neighborhood level, which is presented in column (1) in Table 3. Moreover, as a robustness check, we regress the first difference of the natural logarithm of the number of home appraisals on the first difference of the natural logarithm of unit house prices at the neighborhood level with 100 or more observations, which is presented in column (2) in Table 3. Once again, we find a statistically significant relationship between unit house prices and the number of home appraisals.

Table 3 – Unbalanced Panel Data Analysis at the Neighborhood Level

	(1)	(2)	(3)	(4)	(5)	(6)
	The number of home appraisals			Unit house prices		
Unit house prices	0.455***	0.521***	0.514***			
	(0.0437)	(0.0912)	(0.0497)			
The number of home appraisals				0.0273***	0.0260***	0.0312***
				(0.00268)	(0.00496)	(0.00313)
Constant	0.0312***	0.122***	0.419***	0.0871***	0.0914***	0.0829***
	(0.00800)	(0.0131)	(0.00823)	(0.00168)	(0.00246)	(0.00222)
Year effects	Yes	Yes	Yes	Yes	Yes	Yes
Number of obs.	17,543	6,865	14,307	17,543	6,865	14,307
R-squared	0.423	0.476	0.477	0.076	0.139	0.077
Number of neighborhoods	3,382	1,573	3,358	3,382	1,573	3,358
		100 or more observations	2013-2018		100 or more observations	2013-2018

Robust standard errors in parentheses

*** p<0.01, ** p<0.05, * p<0.1

However, the direction of the relationship might be from unit house prices to the number of home appraisals or it could run both ways. For that reason, we regress the first difference of the natural logarithm of unit house prices on the first difference of the natural logarithm of the number of home appraisals, which is presented in column (4) in Table 3. We observe that the relationship between the number of home appraisals and unit house prices is statistically significant, but the size of the regression coefficient is considerably lower in this case. As another robustness check, we analyze the relationship between the first difference of the natural logarithm of the number of home appraisals

on the first difference of the natural logarithm of unit house prices from 2013 to 2018, which are presented in columns (3) and (6) in Table 3. We observe that the relationship between the number of home appraisals and unit house prices is statistically significant as before. Moreover, the regression coefficients are close to our initial findings.

One of the major issues about administrative units is the partition of old neighborhoods to create new neighborhoods due to rapid population growth and urbanization in Turkey. As a result of that we cannot follow the same neighborhoods over the years. Our unbalanced panel data set collects information from 3,930 neighborhoods from 2010 to 2018, but we can track only 1,203 neighborhoods continuously for the same time period. Moreover, we can track only 2,447 neighborhoods from 2016 to 2018. As a result, we do not perform empirical analysis with a balanced panel data set prepared at the neighborhood level, because its smaller sample size might lead to biased econometric results.

Nevertheless, the presence of data at the neighborhood level gives us the opportunity to follow an alternative approach. We can analyze the direction and the strength of the relationship between unit house prices and the number of home appraisals with respect to the range of available observations (30-99, 100-149, 150-199 and 200 or more). We observe that the relationship between unit house prices and the number of home appraisals becomes stronger as the number of available observations increases steadily, which are presented in columns from (1) to (3) in Table 4. However, the strength of the relationship between unit house prices and the number of home appraisals weakens slightly after the number of available observations pass a certain point, which is shown in column (4) in Table 4. As a result, the econometric results confirm the theoretical proposition that the relationship between unit house prices and the number of home will have a concave shape, which is presented in Figure 2 in Section III.

Table 4 – Unbalanced Panel Data Analysis at the Neighborhood Level

	(1)	(2)	(3)	(4)
	The number of home appraisals			
Unit house prices	0.336***	0.406***	0.495**	0.294**
	(0.0491)	(0.126)	(0.227)	(0.132)
Constant	0.00601	0.160***	0.123***	0.136***
	(0.0102)	(0.0220)	(0.0369)	(0.0197)
Year effects	Yes	Yes	Yes	Yes
Number of obs.	10,581	2,767	1,455	2,740
R-squared	0.349	0.437	0.424	0.479
Number of provinces	2,958	1,262	767	665
	Less than 100 observations	Between 100 and 150 observations	Between 150 and 200 observations	200 or more observations

Robust standard errors in parentheses

*** p<0.01, ** p<0.05, * p<0.1

In a similar fashion, we regress the first difference of the natural logarithm of unit house prices on the first difference of the natural logarithm of the number of home appraisals with respect to the range of available observations (30-99, 100-149, 150-199 and 200 or more). We observe that the relationship between the number of home appraisals and unit house prices strengthens as the number of available observations increases steadily, which are presented in columns from (1) to (3) in Table 5 as before. However, the strength of the relationship between unit house prices and the number of home appraisals weakens slightly after the number of available observations pass a certain point, which is shown in column (4) in Table 5.

Table 5 – Unbalanced Panel Data Analysis at the Neighborhood Level

	(1)	(2)	(3)	(4)
	Unit house prices			
The number of home appraisals	0.0247***	0.0257***	0.0293*	0.0126**
	(0.00367)	(0.00800)	(0.0151)	(0.00604)
Constant	0.0831***	0.0836***	0.0994***	0.100***
	(0.00247)	(0.00509)	(0.00671)	(0.00325)
Year effects	Yes	Yes	Yes	Yes
Number of obs.	10,581	2,767	1,455	2,740
R-squared	0.050	0.092	0.089	0.194
Number of provinces	2,958	1,262	767	665
	Less than 100 observations	Between 100 and 150 observations	Between 150 and 200 observations	200 or more observations

Robust standard errors in parentheses

*** p<0.01, ** p<0.05, * p<0.1

V.2. Provinces

We can analyze 80 provinces from 2010 to 2018, but we do not have observation for each province for every year. We perform fixed effects regressions using our unbalanced panel data set. We also control for year effects in the regressions by introducing time dummy variables. We regress the first difference of the natural logarithm of the number of home appraisals on the first difference of the natural logarithm of unit house prices, which is shown in column (1) in Table 6. We find a statistically significant relationship between unit house prices and the number of home appraisals.

However, as mentioned previously, the direction of the relationship might be from unit house prices to the number of home appraisals or it could run both ways. For this reason, we regress the first difference of the natural logarithm of unit house prices on the first difference of the natural logarithm of the number of home appraisals, which is presented in column (3) in Table 6. We find that there is a statistically significant relationship between the number of home appraisals and unit house prices, but

the size of the regression coefficient is considerably lower in this case. As another robustness check, we regress the natural logarithm of the number of home appraisals on the natural logarithm of unit house prices for provinces from 2013 to 2018, which are presented in columns (2) and (4). We observe a statistically significant relationship between the number of home appraisals and unit house prices as before. Moreover, the explanatory powers of the regressions are slightly higher in these cases.

Table 6 – Unbalanced Panel Data Analysis at the Province Level

	(1)	(2)	(3)	(4)
	The number of home appraisals		Unit house prices	
Unit house prices	1.244**	1.798***		
	-0.503	-0.649		
The number of home appraisals			0.0406***	0.0512***
			-0.0135	-0.015
Constant	-0.0582	0.592***	0.0747***	0.0367***
	-0.0566	-0.0549	-0.00744	-0.0123
Year effects	Yes	Yes	Yes	Yes
Number of obs.	595	457	595	457
R-squared	0.628	0.695	0.142	0.170
Number of provinces	80	80	80	80
		2013-2018 period		2013-2018 period

Robust standard errors in parentheses

*** p<0.01, ** p<0.05, * p<0.1

We can track only 67 provinces continuously from 2010 to 2018.[13] We perform fixed effects regressions using our balanced panel data set. We regress the first difference of the natural logarithm of the number of home appraisals on the first difference of the natural logarithm of unit house prices. We control for year effects in the regressions by introducing time dummy variables. We observe a statistically significant relationship between unit house prices and the number of home appraisals, which is shown in column (1) in Table 7. We also regress the first difference of the natural logarithm of unit house prices on the first difference of the natural logarithm of the number of home appraisals, which is shown in column (3) in Table 7. We observe that there is a statistically significant relationship between the number of home appraisals and unit house prices, but as before the size of the regression coefficient is considerably lower in this case.

As a robustness check, we regress the natural logarithm of the number of home appraisals on the natural logarithm of unit house prices for provinces from 2013 to 2018, which are presented in

[13] The omitted provinces due lack of data are Ağrı (4), Bingöl (12), Bitlis (13), Denizli (20), Gümüşhane (29), Hakkari (30), Muş (49), Tunceli (62), Van (65), Bayburt (69), Şırnak (73), Ardahan (75), Iğdır (76) and Kilis (79), which are small provinces mainly from the Eastern parts of the country.

columns (2) and (4) in Table 7. Once more, we observe a statistically significant relationship between the number of home appraisals and unit house prices. In addition, the explanatory powers of these regressions are higher.

Table 7 – Balanced Panel Data Analysis at the Province Level

	(1)	(2)	(3)	(4)
	The number of home appraisals		Unit house prices	
Unit house prices	0.903*	1.313*		
	(0.533)	(0.665)		
The number of home appraisals			0.0309*	0.0416**
			(0.0155)	(0.0180)
Constant	-0.0319	0.608***	0.0750***	0.0467***
	(0.0555)	(0.0573)	(0.00649)	(0.0138)
Year effects	Yes	Yes	Yes	Yes
Number of obs.	536	402	536	402
R-squared	0.661	0.740	0.142	0.151
Number of provinces	67	67	67	67
		2013-2018 period		2013-2018 period

Robust standard errors in parentheses

*** p<0.01, ** p<0.05, * p<0.1

V.3. Robustness Checks

V.3.a. Unbalanced Panel Data Analysis at the Province Level

First, we perform fixed effects regressions using our unbalanced panel data set. We include provincial population, GDP per capita, occupancy permits and the number of households in a province in the regression of the first difference of the natural logarithm of the number of home appraisals on the first difference of the natural logarithm of unit house prices as control variables.[14] We control for year effects in the regressions by introducing time dummy variables. We introduce the first differences of the natural logarithms of the selected control variables to the regressions as in the case of the main economic variables. We find that there is a statistically significant relationship between the number of home appraisals and unit house prices (Table 8). However, the selected control variables are not statistically significant in the regressions. The estimation periods vary depending on the availability of data.

[14] Occupancy permits are certificates, which are issued for completed or partially completed buildings by the Municipalities for constructions within their boundaries and by Special Administrations such as Organized Industrial Zone Directorates in organized industrial zones and Free Trade Zone Directorates in free zones according to Article 30 of Construction Law No.3194.

Table 8 – Unbalanced Panel Data Analysis at the Province Level with Control Variables

	(1)	(2)	(3)	(4)	(5)
	The number of home appraisals				
Unit house prices	1.243**	1.230**	1.091**	1.792***	1.638**
	(0.505)	(0.513)	(0.541)	(0.646)	(0.723)
Occupancy permits	0.00102				0.0112
	(0.0349)				(0.0447)
Population		0.429			4.891
		(1.618)			(3.354)
GDP per capita			0.516		0.700
			(0.431)		(0.464)
The number of households				2.087	-0.962
				(1.644)	(3.407)
Constant	-0.0584	-0.0583	-0.135	0.529***	0.506***
	(0.0558)	(0.0567)	(0.0856)	(0.0735)	(0.115)
Year effects	Yes	Yes	Yes	Yes	Yes
Number of obs.	595	595	515	457	377
R-squared	0.628	0.628	0.405	0.697	0.439
Number of provinces	80	80	79	80	79
			2010-2017 period	2012-2018 period	2012-2017 period

Robust standard errors in parentheses

*** p<0.01, ** p<0.05, * p<0.1

At the same time, we include provincial population, GDP per capita, occupancy permits and the number of households in a province in the regression of the first difference of the natural logarithm of unit house prices on the first difference of the natural logarithm of the number of home appraisals as control variables. We also control for year effects in the regressions by introducing time dummy variables. We introduce the first differences of the natural logarithms of the selected control variables to the regressions. We find that there is a statistically significant relationship between the number of home appraisals and unit house prices (Table 9). However, the selected control variables are not statistically significant in the regressions and the estimation periods vary depending on the availability of data.

Table 9 – Unbalanced Panel Data Analysis at the Province Level with Control Variables

	(1)	(2)	(3)	(4)	(5)
	Unit house prices				
The number of home appraisals	0.0405***	0.0397***	0.0328**	0.0514***	0.0433***
	(0.0136)	(0.0136)	(0.0135)	(0.0150)	(0.0140)
Occupancy permits	0.00601				0.00828
	(0.00676)				(0.00652)
Population		0.407			0.710
		(0.246)			(0.557)
GDP per capita			-0.0118		-0.0814
			(0.0854)		(0.111)
The number of households				-0.0532	-0.749
				(0.151)	(0.539)
Constant	0.0732***	0.0736***	0.0752***	0.0382***	0.0675***
	(0.00761)	(0.00730)	(0.0148)	(0.0132)	(0.0166)
Year effects	Yes	Yes	Yes	Yes	Yes
Number of obs.	595	595	515	457	377
R-squared	0.143	0.152	0.155	0.170	0.193
Number of provinces	80	80	79	80	79
			2010-2017 period	2012-2018 period	2012-2017 period

Robust standard errors in parentheses

*** p<0.01, ** p<0.05, * p<0.1

V.3.b. Balanced Panel Data Analysis at the Province Level

Second, we perform fixed effects regressions using our balanced panel data set. We include the selected control variables in the regression of the first difference of the natural logarithm of the number of home appraisals on the first difference of the natural logarithm of unit house prices as a robustness check. We find that the regression coefficients of the growth of population and the growth of GDP per capita are statistically significant in the extended model, which is presented in column (5) in Table 10. However, we do not observe a statistically significant relationship between the number of home appraisals and unit house prices, when the growth of GDP per capita is added to the estimation, which is presented in columns (3) and (5). The estimation periods vary depending on the availability of data.

Table 10 – Balanced Panel Data Analysis at the Province Level with Control Variables

	(1)	(2)	(3)	(4)	(5)
	The number of home appraisals				
Unit house prices	0.892*	0.917*	0.714	1.296*	1.066
	(0.534)	(0.538)	(0.473)	(0.662)	(0.658)
Occupancy permits	0.0483				0.0679
	(0.0390)				(0.0471)
Population		-0.921			5.844*
		(1.693)			(3.084)
GDP per capita			0.734*		0.690*
			(0.404)		(0.395)
The number of households				-1.985	-4.866
				(2.320)	(3.861)
Constant	-0.0435	-0.0309	-0.140*	0.668***	0.613***
	(0.0556)	(0.0555)	(0.0725)	(0.0859)	(0.121)
Year effects	Yes	Yes	Yes	Yes	Yes
Number of obs.	536	536	469	402	335
R-squared	0.662	0.662	0.478	0.741	0.527
Number of provinces	67	67	67	67	67
			2010-2017 period	2012-2018 period	2012-2017 period

Robust standard errors in parentheses

*** p<0.01, ** p<0.05, * p<0.1

Moreover, we add provincial population, GDP per capita, occupancy permits and the number of households in a province to the regression of the first difference of the natural logarithm of unit house prices on the first difference of the natural logarithm of the number of home appraisals as control variables. We also control for year effects in the regressions by introducing time dummy variables. We introduce the first differences of the natural logarithms of the selected control variables to the regressions. We find that there is a statistically significant relationship between the number of home appraisals and unit house prices, except when GDP per capita is included in the regression, which is shown in column (3) in Table 11. However, the selected control variables are not statistically significant in the regressions and the estimation periods vary depending on the availability of data. The estimation periods vary depending on the availability of data as before.

Table 11 – Balanced Panel Data Analysis at the Province Level with Control Variables

	(1)	(2)	(3)	(4)	(5)
			Unit house prices		
The number of home appraisals	0.0305*	0.0313*	0.0249	0.0411**	0.0339*
	(0.0156)	(0.0157)	(0.0149)	(0.0180)	(0.0185)
Occupancy permits	0.00487				0.00954
	(0.00807)				(0.00736)
Population		0.269			0.767
		(0.199)			(0.555)
GDP per capita			0.0475		-0.0346
			(0.0863)		(0.123)
The number of households				-0.231	-1.032
				(0.296)	(0.645)
Constant	0.0738***	0.0744***	0.0673***	0.0539***	0.0771***
	(0.00657)	(0.00660)	(0.0144)	(0.0157)	(0.0244)
Year effects	Yes	Yes	Yes	Yes	Yes
Number of obs.	536	536	469	402	335
R-squared	0.143	0.146	0.148	0.152	0.177
Number of provinces	67	67	67	67	67
			2010-2017 period	2012-2018 period	2012-2017 period

Robust standard errors in parentheses

*** p<0.01, ** p<0.05, * p<0.1

VI. Conclusion

We test the validity of the hypothesis that there is a direct relationship between unit house prices and the number of home appraisals rigorously. We find that there is a positive and statistically significant relationship between unit house prices and the number of home appraisals at neighborhood and province levels using both balanced and unbalanced panel data sets. Moreover, the econometric results show that the effect of unit house prices on the number of home appraisals is stronger than that of the number of home appraisals on unit house prices. As a robustness check, we include control variables such as income per capita, housing supply and population growth in the econometric estimations using both balanced and unbalanced panel data sets at province level. The proposition that there is a direct and significant relationship between unit house prices and the number of home appraisals is confirmed by the robustness checks.

Our empirical findings are also supported by aggregate data such as GDP per capita and house sale figures, which show that more house are sold in Western regions and coastal cities, where income

levels are also higher than country average. Income distribution across geographic regions in Turkey, which is measured using micro-economic data from LFS, HBS and SILC, is very close to the distribution of weighted unit house prices that we observe from the RPPI database. Thus, our empirical analysis confirms the hypothesis that there is a direct and significant relationship between unit house prices and the number of home appraisals is valid for the Turkish economy.

The main contribution of this paper is to propose a method to identify wealthy households for oversampling them in household surveys in Turkey. For this purpose, we analyze unit house prices from the CBRT – RPPI database and we reach to the conclusion that unit house prices is a suitable candidate to identify wealthy households and for matching administrative units with the TURKSTAT sampling frame. To the best of our knowledge, an oversampling method based on wealth indicators has not been used before in household surveys in Turkey. As a result, we think that unit house prices in general are a good indicator to identify wealthy households. Different data sources from real estate agencies and internet can also be used to link affluent households and their geographical locations. In the future, a similar approach might be adopted by researchers and also by TURKSTAT for oversampling of wealthy households.

The CBRT – RPPI database represents a high proportion of blocks and households from the NAD across geographic regions, especially in İstanbul. However, the coverage of the CBRT – RPPI database falls considerably, as we proceed from the Western regions towards the Eastern regions. Actually, this is an expected outcome, since the number of house sales is significantly lower in the Eastern provinces. Moreover, the percentage of mortgaged house sales is even lower in these regions, whereas the RPPI database is based on dwelling appraisal reports on house sales, which are subject to mortgage loans. Ideally, the selected database for oversampling should match with the TURKSTAT sampling frame provide at the block level and have a high representation level across all geographical regions of the country.

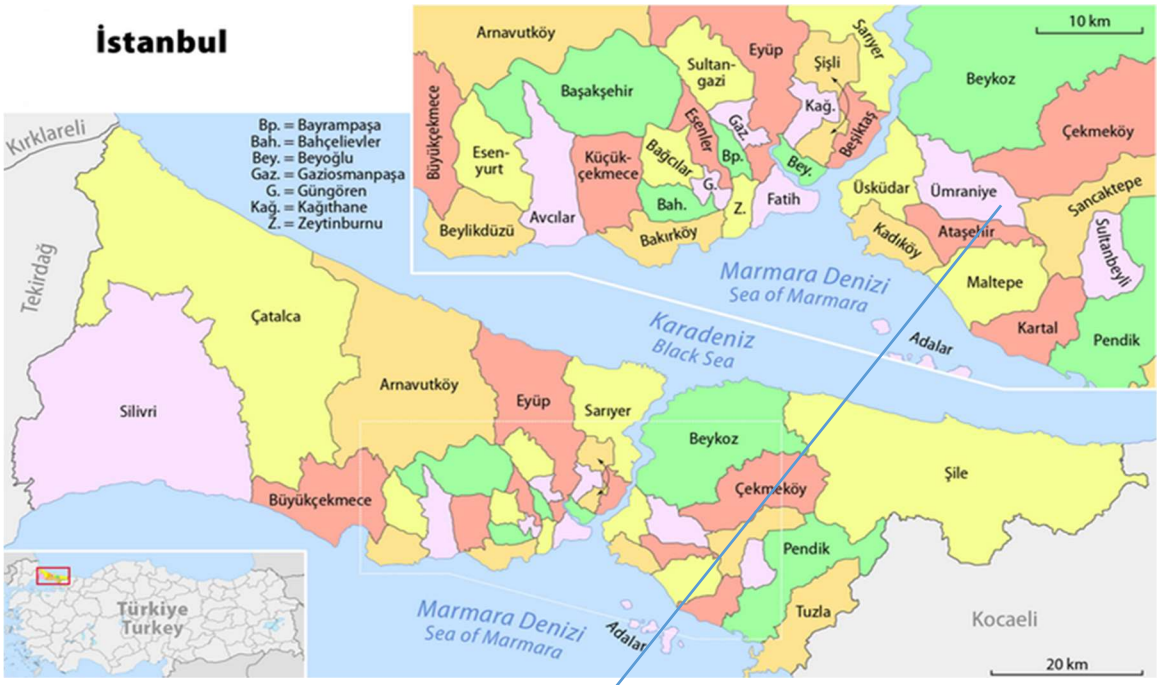
As a caveat, we must mention three important points about the empirical analysis in this paper. First, we must emphasize that our aim is not to try and to establish a structural relationship for the housing market. Second, our identification strategy is based on the analysis of the cross-sectional dimension of the RPPI database. We do not analyze changes in house prices from one year to another. Instead, we link unit house prices and the number of home appraisals at administrative units at each survey year. Finally, the time dimension of our empirical analysis is limited, since the CBRT – RPPI database covers only nine years.

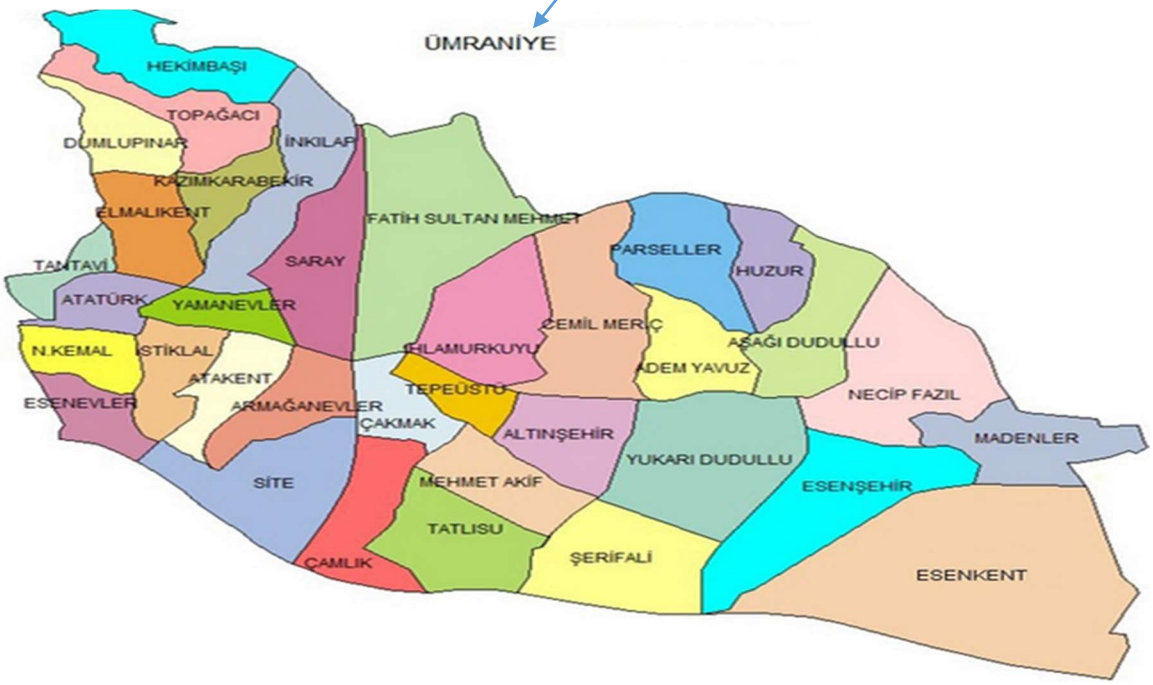
Appendix 1 – Geographical Regions

Table A1.1 – TURKSTAT NUTS1 and NUTS2 Codes

	Regions – NUTS1		NUTS2	Province		Regions – NUTS1		NUTS2	Province
1	TR1	İstanbul	TR10	İstanbul					Zonguldak
2	TR2	West Marmara	TR21	Tekirdağ				TR81	Karabük
				Edirne					Bartın
				Kırklareli				TR82	Kastamonu
			TR22	Balıkesir					Çankırı
				Çanakkale	8	TR8	West Black Sea		Sinop
3	TR3	Aegean	TR31	İzmir				TR83	Samsun
			TR32	Aydın					Tokat
				Denizli					Çorum
				Muğla					Amasya
			TR33	Manisa	9	TR9	East Black Sea	TR90	Trabzon
				Afyonkarahisar					Ordu
				Kütahya					Giresun
				Uşak					Rize
4	TR4	East Marmara	TR41	Bursa					Artvin
				Eskişehir					Gümüşhane
				Bilecik	10	TRA	North East Anatolia	TRA1	Erzurum
			TR42	Kocaeli					Erzincan
				Sakarya					Bayburt
				Düzce				TRA2	Ağrı
				Bolu					Kars
				Yalova					Iğdır
5	TR5	West Anatolia	TR51	Ankara					Ardahan
			TR52	Konya	11	TRB	Middle East Anatolia	TRB1	Malatya
				Karaman					Elazığ
6	TR6	Mediterranean	TR61	Antalya					Bingöl
				Isparta					Tunceli
				Burdur				TRB2	Van
			TR62	Adana					Muş
				Mersin					Bitlis
			TR63	Hatay					Hakkari
				Kahramanmaraş	12	TRC	South East Anatolia	TRC1	Gaziantep
				Osmaniye					Adıyaman
7	TR7	Middle Anatolia	TR71	Kırıkkale					Kilis
				Aksaray				TRC2	Şanlıurfa
				Niğde					Diyarbakır
				Nevşehir				TRC3	Mardin
				Kırşehir					Batman
			TR72	Kayseri					Şırnak
				Sivas					Siirt
				Yozgat					

Appendix 2 – Administrative Units





Appendix 3 – House Sales

Table A3.1 – The Distribution of House Sales across NUTS2 Regions (%)

	2013	2014	2015	2016	2017	2018
Total	100.00	100.00	100.00	100.00	100.00	100.00
TR10	20.29	19.35	18.60	17.33	16.91	17.02
TR21	3.60	3.53	3.60	3.73	3.46	3.33
TR22	2.78	2.89	2.98	3.05	2.93	3.10
TR31	6.26	6.16	6.03	6.06	5.97	5.50
TR32	4.55	4.67	4.62	4.83	5.09	5.10
TR33	2.76	2.82	3.06	3.30	3.55	3.68
TR41	5.67	5.68	5.94	5.92	6.08	5.65
TR42	5.57	5.95	6.13	5.97	6.18	6.25
TR51	11.91	11.31	11.37	10.78	10.68	9.54
TR52	2.60	2.73	2.61	2.62	2.73	2.95
TR61	5.62	5.91	5.59	5.17	4.97	5.35
TR62	4.61	4.47	4.58	4.67	4.82	4.93
TR63	2.58	2.77	2.77	3.05	3.19	3.36
TR71	1.90	1.89	1.81	1.75	1.88	2.00
TR72	3.18	3.20	3.30	3.37	3.26	3.27
TR81	0.83	0.98	1.17	1.11	1.05	1.06
TR82	0.79	0.85	0.88	0.88	0.95	1.03
TR83	3.05	2.96	2.86	3.08	3.12	3.25
TR90	2.56	2.55	2.59	2.61	2.45	2.67
TRA1	0.77	0.80	0.79	0.76	0.84	0.88
TRA2	0.33	0.37	0.41	0.45	0.46	0.48
TRB1	1.90	1.89	1.87	1.98	2.01	1.91
TRB2	0.43	0.40	0.43	1.27	0.80	0.84
TRC1	2.39	2.51	2.44	2.62	2.74	2.83
TRC2	2.27	2.48	2.62	2.61	2.75	2.80
TRC3	0.80	0.89	0.99	1.05	1.13	1.23

Source: TURKSTAT, Authors' calculations

References

Balestra, C and Tonkin, R. (2018), "Inequalities in Household Wealth across OECD Countries: Evidence from the OECD Wealth Distribution Database", OECD Statistics Working Papers No. 2018/01.

Bertaut, C. and Starr-McCluer, M. (2002), *Household Portfolios in the United States*, in Household Portfolios edited by Luigi Guiso, Michael Haliassos and Tullio Jappelli, The MIT Press, Cambridge, MA.

Bricker, J., Henriques, A., Krimmel, J. and Sabelhaus, J. (2016), "Estimating Top Income and Wealth Shares: Sensitivity to Data and Methods", *American Economic Review: Papers & Proceedings*, Vol. 106, No.5, pp. 641-645.

Causa, O., Woloszko, N. and Leite D. (2019), "Housing, Wealth Accumulation and Wealth Distribution: Evidence and Stylized Facts", OECD Economics Department Working Papers No. 1588.

Ceritoğlu, E. (2017), "The Effect of House Price Changes on Cohort Consumption in Turkey", *Central Bank Review*, Vol. 17, No. 3, pp. 99-110.

- (2020), "Home-ownership, Housing Demand, and Household Wealth Distribution in Turkey", *Emerging Markets Finance and Trade*, Vol. 56, No. 5, pp. 1146-1165.

Chakraborty, R. and Waltl, S.R. (2018), "Missing the Wealthy in the HFCS: Micro Problems with Macro Implications", ECB Working Paper No. 2163.

Chakraborty, R., Kavonius, I.K., Pérez-Duarte, S. and Vermeulen, P. (2019), "Is the Top Tail of the Wealth Distribution the Missing Link between the Household Finance and Consumption Survey and National Accounts?", *Journal of Official Statistics*, Vol. 35, No. 1, pp. 31-65.

Goodman, A.C. (1988), "An Econometric Model of Housing Price, Permanent Income, Tenure Choice and Housing Demand", *Journal of Urban Economics*, Vol. 23, No.3, pp. 327-353.

- (1990), "Demographics of Individual Housing Demand", *Regional Science and Urban Economics*, Vol. 20, No.1, pp. 83-102.

Halicioglu, F. (2007), "The Demand for New Housing in Turkey: An Application of ARDL Model", *Global Business and Economics Review*, Vol. 9, No. 1, pp. 62-74.

Household Finance and Consumption Network (2016a), "The Household Finance and Consumption Survey: Methodological Report for the Second Wave", ECB Statistics Paper Series No. 17.

- (2016b), "The Household Finance and Consumption Survey: Results from the Second Wave", ECB Statistical Paper Series No. 18.

Hülagü, T., Kızılkaya, E., Özbekler, A.G. and Tunar, P. (2016), "A Hedonic House Price Index for Turkey", CBRT Working Papers, No. 16/03.

Kalton, G. (2009), "Methods for Oversampling Rare Subpopulations in Social Surveys", *Survey Methodology*, Vol. 35, No. 2, pp. 125-141.

Kennickell, A.B. (2008), "The Role of Over-sampling of the Wealthy in the Survey of Consumer Finances", *Irving Fisher Committee Bulletin*, Vol. 28, pp. 403-408.

Kiel, K.A. and Zabel, J.E. (2008), "Location, location, location: The 3L approach to house price determination", *Journal of Housing Economics*, Vol. 17, No. 2, pp. 175-190.

Knoll, K., Schularick, M. and Steger, T. (2017), "No Price Like Home: Global House Prices, 1870-2012", *American Economic Review*, Vol. 107, No. 2, pp. 331-353.

Marpsat, M. and Razafindratsima, N. (2010), "Survey Methods for Hard-to-Reach Populations: Introduction to the Special Issue", *Methodological Innovations Online*, Vol. 5, No. 2, pp. 3-16.

Riddel, M. (2004), "Housing-Market Disequilibrium: An Examination of Housing-Market Price and Stock Dynamics 1967–1998", *Journal of Housing Economics*, Vol. 13, No. 2, pp. 120-135.

Rosen, K.T. and Smith, L.B. (1983), "The Price-Adjustment Process for Rental Housing and the Natural Vacancy Rate", *American Economic Review*, Vol. 73, No. 4, pp. 779-786.

Steiner, E. (2010), "Estimating a Stock-Flow Model for the Swiss Housing Market", *Swiss Journal of Economics and Statistics*, Vol. 146, No. 3, pp. 601-627.

Valliant, R., Hubbard, F., Lee, S. and Chang, C. (2014), "Efficient Use of Commercial Lists in U.S. Household Sampling", *Journal of Survey Statistics and Methodology*, Vol. 2, pp. 182–209.

Vermeulen, P. (2016), "Estimating the Top Tail of the Wealth Distribution", *American Economic Review: Papers & Proceedings*, Vol. 106, No. 5, pp. 646-650.

- (2018), "How Fat is the Top Tail of the Wealth Distribution?", *Review of Income and Wealth*, Vol. 64, No. 2, pp. 357-387.

Zabel, J.E. (2004), "The Demand for Housing Services", *Journal of Housing Economics*, Vol. 13, No. 1, pp. 16-35.

Central Bank of the Republic of Turkey

Recent Working Papers

The complete list of Working Paper series can be found at Bank's website

(**http://www.tcmb.gov.tr**)

Corporate Debt Maturity, Repayment Structure and Monetary Policy Transmission
(Hatice Gökçe Karasoy Can Working Paper No. 20/09, May 2020)

Detecting the Position of Countries in Global Value Chains Using a Bilateral Approach
(Oğuzhan Erdoğan Working Paper No. 20/08, May 2020)

A Measure of Turkey's Sovereign and Banking Sector Credit Risk: Asset Swap Spreads
(Abdullah Kazdal, Halil İbrahim Korkmaz, Doruk Küçüksaraç, Yiğit Onay Working Paper No. 20/07, May 2020)

Nowcasting Turkish GDP Growth with Targeted Predictors: Fill in the Blanks
(Mahmut Günay Working Paper No. 20/06, May 2020)

Bank Lending and Maturity: the Anatomy of the Transmission of Monetary Policy
(Selva Bahar Baziki, Tanju Çapacıoğlu Working Paper No. 20/05, March 2020)

Do Local and Global Factors Impact the Emerging Markets's Sovereign Yield Curves? Evidence from a Data-Rich Environment
(Oğuzhan Çepni, İbrahim Ethem Güney, Doruk Küçüksaraç, Muhammed Hasan Yılmaz Working Paper No. 20/04, March 2020)

The Role of Imported Inputs in Pass-through Dynamics
(Dilara Ertuğ, Pınar Özlü, M. Utku Özmen, Çağlar Yüncüler Working Paper No. 20/03, February 2020)

Nowcasting Turkish GDP with MIDAS: Role of Functional Form of the Lag Polynomial
(Mahmut Günay Working Paper No. 20/02, February 2020)

How Do Credits Dollarize? The Role of Firm's Natural Hedges, Banks' Core and Non-Core Liabilities
(Fatih Yılmaz Working Paper No. 20/01, February 2020)

Hidden Reserves as an Alternative Channel of Firm Finance in a Major Developing Economy
(İbrahim Yarba Working Paper No. 19/36, December 2019)

Interaction of Monetary and Fiscal Policies in Turkey
(Tayyar Büyükbaşaran, Cem Çebi, Erdal Yılmaz Working Paper No. 19/35, December 2019)

Cyclically Adjusted Current Account Balance of Turkey
(Okan Eren, Gülnihal Tüzün Working Paper No. 19/34, December 2019)

Term Premium in Turkish Lira Interest Rates
(Halil İbrahim Aydın, Özgür Özel Working Paper No. 19/33, December 2019)

Decomposing Uncertainty in Turkey into Its Determinants
(Emine Meltem Baştan, Ümit Özlale Working Paper No. 19/32, December 2019)

Demographic Transition and Inflation in Emerging Economies
(M. Koray Kalafatcılar, M. Utku Özmen Working Paper No. 19/31, December 2019)

Facts on Business Dynamism in Turkey
(Ufuk Akçiğit, Yusuf Emre Akgündüz, Seyit Mümin Cılasun, Elif Özcan Tok, Fatih Yılmaz Working Paper No. 19/30, September 2019)

Monitoring and Forecasting Cyclical Dynamics in Bank Credits: Evidence from Turkish Banking Sector
(Mehmet Selman Çolak, İbrahim Ethem Güney, Ahmet Şenol, Muhammed Hasan Yılmaz Working Paper No. 19/29, September 2019)

Intraday Volume-Volatility Nexus in the FX Markets: Evidence from an Emerging Market
(Süleyman Serdengeçti, Ahmet Şensoy Working Paper No. 19/28, September 2019)

Is There Asymmetry between GDP and Labor Market Variables in Turkey under Okun's Law?
(Evren Erdoğan Coşar, Ayşe Arzu Yavuz Working Paper No. 19/27, September 2019)