Heterogeneities in the House Price Elasticity of Consumption

Nazanin Khazra*

November 14, 2019

Latest version available here

Abstract

I provide new evidence on the house price elasticity of consumption by exploiting micro-level consumption data from the Nielsen consumer panel for 2004 to 2016. I estimate the elasticity as a non-parametric function of household characteristics using a newly developed causal machine learning model called Generalized Random Forest (GRF). I find substantial county and household level heterogeneities in the elasticity. At the county level, it ranges from 0.04 to 0.16 with some neighboring counties being up to eight standard deviations apart. On the household level, elasticity ranges from 0.01 to 0.21 in which household structure plays an important role in defining the heterogeneity. Among all characteristics, having a child, size of the household, and age of the heads of the household create substantial disparities. Therefore, basing policies on an average estimate and avoiding local and household level heterogeneities may result in unintended policy consequences. I find that locations with volatile housing markets are less elastic; thus, not accounting for local heterogeneities overestimates total consumption responses in booms and underestimates them in busts. This is the first paper that studies the heterogeneities in the house price elasticity of consumption at the household level and highlights the importance of regional and time variations in this elasticity.

*Department of Economics, University of Illinois at Urbana-Champaign, 214 David Kinley Hall, 1407 West Gregory Drive, Urbana, IL 61801, USA. Email:khazra2@illinois.edu.
I am extremely greatful to Prof. Dan Bernhardt, Prof. David Albouy, and Prof. Dan McMillen for their guidance and advice during this research and to Prof. Alex Bartik and Prof. Greg Howard for their endless support and supervision. I am also thankful to Prof. Mark Borgschulte, Prof. Ben Marx, Prof. Andrew Garin, and Prof. Elizabeth Forsyth for their valuable comments and discussions.

1 Introduction

Housing is a large part of an average American household's wealth and highly volatile in value. Hence, it is important to know the causal effect of house prices on consumption. This relationship gained more attention after the Great Recession. A wide body of literature following a series of Atif Mian, Amir Sufi and coauthors papers[1] has shown that on average households respond to positive house price shocks by increasing consumption. However, basing policies on an average estimate for the entire country introduces biases, and these biases are especially large during busts. House price elasticity of consumption varies regionally and even though the average may remain constant over time, the local heterogeneities are considerably large.

Despite the consensus on the existence and relative magnitude of the average causal effect of house price (or wealth) shocks on consumption, there are important unanswered questions in this literature. How representative is the average estimate of elasticity for different localities? What does the distribution of regional elasticities look like? Are there boom-bust asymmetries? Who are the people that are more sensitive to house price shocks? To what extent is the house price elasticity of consumption expenditures heterogeneous? What are the underlying reasons for this heterogeneity in terms of fundamental variables and household characteristics? Answering these questions has important implications for modeling the macroeconomy and for policy making.

This is the first paper that estimates the house price elasticity as a function of a wide set of household characteristics together with time and location and shows a full picture of the important dimensions. I use a newly developed causal Machine Learning (ML) model called Generalized Random Forest (GRF) which was introduced in Athey et al. (2019b) to estimate this function. Using the estimated relation, I predict the elasticity at the household and county levels over time. I then show how this elasticity varies by household characteristics and by counties. Finally I show that disregarding the local heterogeneities can result in aggregation bias and mis-measurements in consumption changes.

There are some existing papers that have emphasized the importance of regional heterogeneity in income, leverage, and wealth. Using a cross sectional IV regression, Mian et al. (2013) find large heterogeneity across zipcodes by income and leverage. In their study the correlation between other characteristics of the zipcode and income such as age is not considered in the regressions. In my paper, I take the correlations into account by estimating the elasticity as a function of a wide set of controls. Cloyne et al. (2019) study the effect of house prices on household borrowing and control for the heterogeneity in age, income, income growth and loan-to-value dimension and find no

[1] A small subset of references are Campbell and Cocco (2007), Mian and Sufi (2011), Mian et al. (2013), Mian and Sufi (2014), Mian and Sufi (2016), Stroebel and Vavra (2016), Aladangady (2017), Guren et al. (2018) , Graham (2018), and Aruoba et al. (2019).

heterogeneity except for loan-to-value. They estimate the heterogeneities using a binned regression approach. Unlike their paper, I do not have any functional form assumptions on the heterogeneities. And finally, Beraja et al. (2018) show that time varying regional distribution of wealth matters for the aggregate effect of monetary stimulus policies and for ensuring that the stimulus flows to the regions that need it most. In line with their paper, my estimations of the regional heterogeneities in the house price elasticity have implications for policymakers to target the right populations.

I show that disregarding the local heterogeneities leads to overestimation in consumption responses during booms and underestimation during busts (recession years). The local heterogeneities in house price elasticity matter because the covariance between house price changes and the elasticity is not zero. It is positive for booms and negative for busts. The patterns in the covariance is mainly coming from the dispersion between locations with high and low volatility in the housing market. Counties with high volatility in house prices are on average less elastic but have larger house price changes. Treating these locations as average results in measurement biases in the aggregate consumption responses.

My results have applications in two main areas. First, I estimate the elasticity as a function of household characteristics, region and time. This has implications for policymakers and allows them to predict the elasticity over time and regional dimensions. Being able to estimate the elasticity for different regions and over time can be of use to policymakers to have more precise estimations of the aggregate effects of house prices on consumption and to design effective and efficient policies particularly during recessions. Second, the representative household looks different in different regions and over time. My results show that household characteristics such as having a child, size of the household, and age of the heads of the household create large disparities between households in terms of their elasticity. These dimensions of heterogeneities are data driven and can help make macro models to be closer to the reality by considering these facts in modeling preferences and constraints.

I use Kilts-Nielsen Consumer Panel Dataset (KNCP) for non-durable expenditures data. This is a panel dataset of about 40,000 to 60,000 households' daily supermarket and hypermarket shopping trips. For house prices, I use house price indexes (HPI) from the Federal Housing Finance Agency. Using household level consumption data I estimate the house price elasticity at the individual level and look at heterogeneities in a more comprehensive way that can shed more light on the underlying mechanisms.

I use GRF to estimate a non-parametric relation between the house price elasticity of consumption and household characteristics. From the ML perspective, this method is an extension of the Random Forest (RF) model which is getting more popular in applied fields [2]. Both GRF and RF are ensemble tree-based ML models which base the

[2]Some recent papers that use GRF are Davis and Heller (2017), Hoffman and Mast (2019), Carter et al.

estimation on recursive partitioning, sub-sampling and random split selection. In contrast to the traditional ML methods such as RF which minimize the error of prediction, GRF is built to find the causal relation between the outcome and a potentially endogenous variable. This method minimizes weighted moment conditions where weights are derived using RF. This method is similar to traditional kernel-based regression models. The difference is that the weights are not constant and the model estimates a data driven measure of neighborhoods and weights for each observation.

In my estimation, I contrast the results for both linear instrumental variable (IV) regression and GRF model. In the first section of the paper, I estimate the conventional IV regression of consumption expenditures growth on house price growth. I instrument for house price growth using local house price sensitivity to regional house price cycles estimations by Guren et al. (2018). They estimate a CBSA level sensitivity instrument using data for pre-1990 and control for a comprehensive set of fixed effects and local economy variables such as employment. The identification assumption is that historic house price sensitivity is correlated with changes in consumption only through changes in house prices. Using this instrument and a linear IV regression, I find an elasticity of 0.09.

Next, using GRF models and same control variables and fixed effects as in the linear IV regression, I estimate the distribution of the elasticity as a function of different household characteristics. The average elasticity using the GRF model is 0.083 which is close to the linear IV regression results. On average a 1% change in real HPI growth results in 0.083% of change in real consumption expenditures growth. Among all, household age, having a child, and size are important for the level of response. My results confirm the literature findings on age, I find that younger households are 37% more elastic than older households by 0.027. A finding new to the literature is that having a child is almost as important as age in shaping the house price response of households. Controlling for all other characteristics, having children increases the elasticity of households by 0.025 which is 35% of the elasticities of the households without a child. Results show that the observed effect of having a child on elasticity is not attributed to marital status or size of the household.

The rest of the paper is structured as the following. Section 2 explains the data used, section 3 provides the linear IV regression results. Next, in section 4, I explain the Generalized Random Forest model, section 5 provides the heterogeneity results and contrasts it with the findings of the literature. Section 6 discusses the policy implications and section 7 describes the aggregation bias and boom-bust asymmetries. Section 8 concludes.

(2019), Daoud and Johansson (2019), Cui and Davis (2019) and Burauel and Schroeder (2019).

2 Data and Descriptive Model

2.1 Nielsen Consumer Panel

The non-durable expenditures data comes from the Kilts-Nielsen Consumer Panel (KNCP)[3]. This dataset is a daily panel of about 40,000 to 60,000 households for the 2004 to 2016 period and consists of quantity and price information on supermarket and hypermarket purchases such as food, cleaning supplies, health and beauty care and some general merchandise that can include more durable goods such as office supplies and electronics. After each shopping trip, households record their total expenditures.

Each household reports some demographic information to the Nielsen Center and updates this information annually. These demographics include household size, income (in 16 bins)[4], age, presence and age of children, employment status, occupation, weekly work hours, educational attainment for male and female household heads, marital status, type of residence and race. The summary statistics of the household demographics is provided in table 1.

On average, households remain in the panel for about 4 years and 13% of the households remain for the entire sample period (13 years)[5]. Households also report their zipcode of residence. At last, Nielsen provides survey weights to adjust for the higher probability of some households being selected and also make the sample to be representative of the U.S in many dimensions such as household size, income, age and education of household head, presence of children and county size. Therefore, regressions in this paper are weighted by these survey weights.

I then aggregate daily households expenditures to an annual level since it is unlikely for any of the underlying channels (collateral or wealth) to operate immediately after an increase in house prices. The causal effect of house prices can mostly show in a more aggregated level especially for highly non-durable goods (such as food). Even in a durable products context, households still need some time to practically respond to the house price changes (it will take approximately 2 to 4 weeks for a home equity line of credit to be processed)[6].

I drop observations with monthly total expenditure to income ratio of higher than 20% [7] or with monthly expenditure of larger than \$3,000[8]. I normalize income and

[3]The panel data is provided by Nielsen through the Kilts Center.

[4]Panelists report their income the fall prior to the panel year.

[5]After winsorizing and dropping the inconsistent reports, I have 7,909 households that participate in the survey for 13 years

[6]One can apply online for such loans and they can request for as low as \$25 and as high as about 85% of the value of the house.

[7]Since consumption of these households are quite persistent, my conjecture is that they made an honest mistake and reported monthly income instead of annual.

[8]In my sample there are about 1,653 households that meet either of these properties

consumption expenditures by Consumer Price Index (CPI)[9] of 2004. In my sample, on average households spend about \$6,000 on groceries in a given year, presented in table 2. The household heads are on average about 56 years old and they have an average income of about \$51,000. After dropping the outliers and subsequent merges, I have 335,508 observations of about 34,000 households every year and these households cover about 8,000 zipcodes and 900 counties.

2.2 House Price Data

I use the publicly available house price index from the Federal Housing Finance Agency (FHFA). The index is a repeat sales (or refinancing) measure for single family house prices calculated based on mortgages purchased or securitized by Fannie Mae or Freddie Mac. The HPI index that I use is at the annual and CBSA[10] level and is not seasonally adjusted.

Average HPI growth for counties is -1% over time and in my sample as in table 2. Growth in house prices has large variation over time and locations. Table 3 shows the counties with the highest house price growth among all counties in the U.S. in each year. In 2017, Holt county in Missouri experienced a 39% growth in HPI from 2016. In 2007, the first year of recession, house prices in Love county Oklahoma increased by 20% despite the huge downturn in the housing market. This may be due to the oil and gas industry in this county. In addition to the county level HPI index, I also use the national level estimates of the house price index from FHFA and normalize it with the CPI of 2004 to create the instrument which I will explain in more details in section 3.1.

2.3 Descriptive Model

This section uses a simplified version of Berger et al. (2017)'s model to illustrate the possible channels through which house prices can affect consumption. It is worth mentioning that this paper does not intend to disentangle and estimate these channels. The purpose of presenting this model is to clearly explain the two main causal channels for the relationship between house prices and consumption discussed in the literature.

Consider a simple model where households (i) live for J periods and hold a risk-free asset (A_{it}) and housing (H_{it}). Housing is traded at P_t and provides one-for-one housing services. Households control variables are current period's consumption (C_{it}) and their stock of the riskless asset (A_{it+1}) for the next period. Households can borrow at most

[9] CPI for Urban areas, not seasonally adjusted, base year of 2004 from U.S. Bureau of Labor Statistics (2018).

[10] A core-based statistical area (CBSA) is defined by the United States Office of Management and Budget (OMB) as metropolitan and micropolitan areas. Metropolitan and micropolitan statistical areas respectively have at least one urbanized area of above 50,000 and 10,000 population.

$(1-\theta)$ of the discounted value of their house which determines the collateral constraint (equation 2).

Income has an exogenous process of the form $Y_{it} = exp(\chi_{it}) + z_{it}$, where z_{it} is a stationary shock with an AR(1) process of the form $z_{it} = \rho_{it-1} + \epsilon_{it}$. With this set up, the maximization problem for a representative household is the following bellman equation:

$$V(P_t, z_{it}) = \underset{C_{it}, A_{it}}{max} U(C_{it}, H_{it}) + \beta E\big[V(P_{t+1}, z_{it+1})\big] \tag{1}$$

subject to

$$\begin{aligned} &C_{it} + A_{it+1} = Y_{it} + (1+r)A_{it} \\ &-A_{it+1} \leq \frac{1-\theta}{1+r} P_t H_{it} \\ &Y_{it} = Y_{it-1} + z_{it} \end{aligned} \tag{2}$$

Considering this simple maximization problem. House price increases can affect consumption through two channels, the overall increase in wealth, the "wealth effect", and the looser borrowing constraint, the "collateral effect".

The wealth effect directly refers to the Permanent Income Hypothesis (PIH) notion of the permanent income/wealth increase. If we consider housing an asset, then its appreciation is analogous to an increase in wealth which consequently results in an increase in consumption. There are mixed findings on the wealth effect channel. For instance, by focusing on households with high credit scores (who are not considered financially constraint), Aruoba et al. (2019) find no effect of house prices on consumption which they interpret as no wealth effect. On the other hand, Campbell and Cocco (2007) find a positive wealth effect.

The collateral channel has been the most agreed upon mechanism in the empirical literature. Many papers have successfully found evidence showing that households in practice do use their appreciated housing wealth as collateral for consumption[11]. The following sections provide the empirical methodologies to estimate the overall causal relationship between house prices and consumption expenditures and discuss the results.

3 Linear Instrumental Variable Regressions

In order to estimate the house price elasticity of non-durable consumption expenditures, I follow the literature and run the regression below.

$$\Delta c_{it} = \beta_1 \Delta hp_{ct} + \beta_2 X_{it} + \alpha_t + \alpha_{ts} + \alpha_{cbsa} + \epsilon_{it} \tag{3}$$

[11]Examples are Aladangady (2017), Campbell and Cocco (2007), Wong (2019).

As in Aladangady (2017)'s notation, lower case c and hp respectively correspond to log changes in real consumption expenditures of each household and the house price index of the county[12]. X_{it} is household level controls such as type of residence, time varying household income, average age, age squared, education, employment and occupation of heads of the household, household size, marital status and race. I include year (), CBSA (α_{cbsa}) and year-state (α_{ts}) fixed effects to control for location and time specific trends of both expenditures and house prices.

The coefficient of interest (elasticity) is β_2 which can be interpreted as the percentage of change in a household's consumption expenditure as a response to a 1% change in house prices. There exists strong co-movements between consumption expenditures and house prices as shown by the OLS estimates in table 5. The correlation between house price growth and consumption expenditures growth is about 0.05 and significant. However, other factors such as unobservable time varying local economy conditions, households' future expectations of their income and the state of the economy can endogenously affect both house prices and consumption. One possible unobservable may be the opening of a new supermarket in the zipcode which can increase both house prices through amenities or gentrification (look at Pope and Pope (2015) and Sullivan (2014)) and consumption due to easier access[13].

Reverse causality between consumption expenditures and house prices can also arise through general equilibrium. Higher consumption can improve local businesses and increase house prices and therefore bias the estimations upward. Moreover, there may be slight spillovers through the labor market. For instance, in the case of a positive exogenous shock to the local housing market, individuals who are working in the construction sector can experience a positive income shock and consequently increase their consumption.

To address the possible endogeneity, reverse causality and possible measurement error concerns, I use an instrumental variable approach from Guren et al. (2018). In short, their instrument takes advantage of the "systematic differences in the sensitivity of local house prices to aggregate shocks across CBSAs". The following section describes the instrument and the identification assumption.

[12]
$$\Delta c_{ict} = log(C_{ict}) - log(C_{ict-1})$$
and similarly for house prices
$$\Delta hp_{ict} = log(HP_{ict}) - log(HP_{ict-1})$$

[13]Using a DID strategy Pope and Pope (2015) find that opening a Walmart increases housing prices by 2% to 3% for nearby houses. In another study, Sullivan (2014) show that organic supermarkets and the ones that promote healthy living cause gentrification. Their finding introduces another channel from which house prices can be effected.

3.1 Instrumental Variable

I use Guren et al. (2018)'s proposed instrumental approach to instrument for the change in house prices. They introduce an instrument that uses systematic differences in CBSA-level exposure to regional house price cycles as an exogenous source of variation in local house prices.

The instrument is built based on two facts; first, the regional house price cycles exist[14]. These cycles happen in different times and most of the time they are averaged out at the national level. Therefore, one can distinguish between the regional and national house price cycles (time) and also between different regions (geography). The second fact is that the sensitivity patterns look stable. In other words, more sensitive areas stay sensitive over time. This fact is central to the identification assumption in that it shows that the estimated sensitivities are exogenous to time varying local economy variables after controlling for location fixed effects.

Guren et al. (2018) estimate their CBSA level sensitivity measure ($\hat{\gamma}$) as the regression coefficient of the log difference of CBSA house prices on log difference of region house prices controlling for changes in CBSA and region employments, region trends and an exhaustive set of local economy variables[15]. This instrument allows for CBSA level heterogeneity that merely comes form the "residual sensitivity" of house prices conditional on the local economy conditions. The $\hat{\gamma}$s are estimated using data for pre-1990 and are not correlated with the current local economy variables conditional on the controls and fixed effects.

The variations in CBSA sensitivities are likely driven by the supply constraints. Guren et al. (2018) interpret the sensitivity estimates as the *"cross-sectional variation in the slope of current or perceived future housing supply curves across CBSAs"*. Given a positive shock to the regional house prices, more sensitive (less supply elastic) locations experience a higher price increase. The supply constraints captured in the estimated sensitivities are more realistic compared to the land unavailability measures by Saiz (2010)[16]. The

[14]Theoretically regional cycles could be associated to households' expectations. Using a reverse engineering approach and an asset pricing model, Gelain et al. (2015) show that an assumption to replicate the U.S. housing market's boom-bust cycles is that households have a "simple random walk forecasting rule" for house prices.

[15]Their regression specifically is the following

$$\Delta HP_{CBSA*time} = \phi_{CBSA} + \delta_{CBSA}\Delta Empl_{CBSA*time} + \mu_{CBSA}\Delta Empl_{region*time}$$

$$+\gamma_{CBSA}\Delta HP_{region*time} + controls + \epsilon$$

Where $Empl$ is the retail employment per capita, and X contains 30-year fixed mortgage rate or a measure of excess bond premium, changes in average wage and some 2 digit industry shares. Moreover, they use the leave-one-out approach for the construction of the regional house price changes which means that $CBSA_i$'s is left out when estimating the sensitivity for that CBSA.

[16]A majority of the papers in this literature use supply elasticity measures introduced in Saiz (2010) based on topological measures for land unavailability and Guren et al. (2018) based on local housing supply regulations as instruments for growth in house prices.

sensitivity measure contain information on the already existing housing stock beside the geographic restrictions measured in land unavailability measures.

The sensitivity instrument is similar to Palmer (2015)'s variability instrument. Palmer (2015) interacts the variance of historic (1980-1995) monthly regional house prices by the changes in national house prices. The CBSA house price variances, however, is likely to be correlated with local economy characteristics such as the share of different industries from the local economy. On the other hand, the sensitivity IV partializes the local economy characteristics by controlling for a rich set of controls including changes in the retail employment at the local and regional level and CBSA fixed effects.

I therefore instrument for changes in local house prices by the interaction of Guren et al. (2018)'s sensitivity estimates ($\hat{\gamma}_{CBSA}$) and log changes in national house prices similar to their paper, expressed in the following formula:

$$Z_{CBSA,t} = \hat{\gamma}_{CBSA} \Delta hp_{U.S.,t} \tag{4}$$

Interacting the sensitivity estimates with national house price trends allows for both time and regional variations in the instrument which will consequently help with identifying the house price elasticity over time.

3.1.1 Comparison with the Supply Elasticity IV

As mentioned before, a majority of the causal papers that have estimated a measure of housing wealth or house price elasticity of consumption use the supply elasticity measures from Saiz (2010) and/or Wharton Residential Land Use Regulation Index (WRLURI) from Gyourko et al. (2008). In the next few paragraphs, I provide a brief explanation of the supply elasticity indexes and compare them to Guren et al. (2018)'s price sensitivity instrument.

In a nutshell, Saiz creates an index for the ease of construction mainly based on land unavailability (estimated using GIS maps). This measure is non-linearly higher for flatter metropolitan areas with less bodies of water that have higher supply elasticity. On the contrary, steeper topology with more lakes and oceans are considered low elasticity MSAs.

To build WRLURI, Gyourko et al. (2008) create an index for the institutional difficulty (state and local stringency) of building new housing. They use a survey of over 2000 jurisdictions in the US that documents the regulations and other local policies surrounding new housing such as land use regulations. These indexes are claimed to capture exogenous variations in local housing supply that are correlated with house prices but not with local economy shocks and other endogenous determinants of consumption.

The local house price sensitivity IV has two main advantages over the Saiz supply elasticity measure and WRLURI. One, it is more comprehensive geographically and two, it is stronger in that it is incorporating more determinants of housing supply rather than just land availability and regulations. Furthermore, there are some recent concerns in using these instruments as exogenous determinants of local house prices which makes the local house price sensitivity IV a more appropriate choice. Guren et al. (2018) estimate this measure for 380 CBSAs that are geographically smaller than the level of Saiz's supply elasticity measure (metropolitan areas). Besides, their measure covers a more representative geography compared to Saiz's. Second, Guren et al. (2018)'s sensitivity measure is estimated conditional on the local economy time varying characteristics and region trends which addresses the main critique to the Saiz supply elasticity measure, endogeneity to the industry composition.

Howard and Liebersohn (2018) and before them Davidoff et al. (2016) bring up the possible correlation between demand shocks and housing supply constraints. For instance, Howard and Liebersohn (2018) model the possible relationship between industry composition and the elasticity of supply for housing. They argue that in places where the supply elasticity is higher, manufacturing industries are more pronounced and this phenomenon will amplify the effect of economy wide shocks on house prices. If there exists such correlation, it may contradict the identification assumption of the supply shifter IV. Another critique is that high skilled workers tend to sort into areas closer to oceans. Davidoff et al. (2016) argues that although oceans are not correlated with demand shocks per se, high skilled workers selection into coastal areas will change the composition of industries and work force and eventually creates the correlation with income distribution and demand shocks.

Both the sensitivity estimate and Saiz measure show similar more inelastic housing supply patterns for coastal regions and more elastic or sensitive patterns for the regions in the interior of the country. Besides the general similarities between the two measures, the rankings of the sub-regions (CBSA in Guren et al. (2018) measure and MSA in Saiz (2010)) are quite different.

As Guren et al. (2018) mention, the correlation between their measure and Saiz is only 15%. One reason for this difference could be that Guren et al. (2018)'s measure considers future housing market elasticity as well as current. They provide the following example about the difference between Pittsburgh and Las Vegas. These two cities have similar land unavailability measures but Guren et al. (2018)'s sensitivity measure predicts a higher value for Las Vegas compared to Pittsburgh. They argue that the better future prospects in Las Vegas (or in other words the optimism) is responsible for the difference between the two cities.

Another important housing market feature that is captured in Guren et al. (2018)'s IV but not in Saiz is that houses are durable goods and therefore, the supply elasticity is low in downturns and also the following upturns. A good example is Detroit. Saiz elasticity for Detroit is 1.24 which is about 0.7 standard deviation smaller than the median elasticity in their sample of MSAs. However, Guren et al. (2018)'s sensitivity estimate is 1.55 which is 1.8 standard deviation higher than the median sensitivity subsequently the sensitivity measure defines Detroit as a relatively inelastic city whereas Saiz's measure categorizes Detroit as an elastic city.

In short, Guren et al. (2018)'s instrument as opposed to Saiz's elasticity measure, captures more of the local characteristics of the housing market such as the previous and future trends (bubbles) of the city and also is more geographically comprehensive.

3.2 IV Regression Model and the Curse of Dimensionality

I instrument for changes in the local house prices by the interaction of Guren et al. (2018)'s sensitivity estimates ($\hat{\gamma}_{CBSA}$) and log changes in national house prices similar to their paper, expressed in equation 4. The identification assumption is that there is no unobserved variable (that is correlated with national HPI changes) that differentially impacts changes in consumption and is correlated with $\hat{\gamma}$ after controlling for local time varying observables and time trends.

The first stage results are presented in table 4. The IV is strongly predicting the house price changes and the projected F-statistics are very strong in all of the specifications[17]. The following are the two stages of the IV regressions.

$$\begin{aligned}
\Delta hp_{ct} &= \beta_1' \hat{\gamma}_{CBSA} \Delta hp_{U.S,t} + \beta_2' X_{it} + \alpha_t + \alpha_{ts} + \alpha_{cbsa} + e_{ict} \\
\Delta c_{it} &= \beta_1 \widehat{\Delta hp_{ct}} + \beta_2 X_{it} + \alpha_t + \alpha_{ts} + \alpha_{cbsa} + \epsilon_{it}
\end{aligned} \tag{5}$$

Columns 3 and 4 of table 5 shows the IV estimates for house price elasticity of consumption. The elasticity is estimated to be 0.09 which means that a 1% increase in house prices increases the non-durable consumption expenditures by 0.09%. This finding is in range with the literature's finding on the house price elasticity of non-durable consumption. Section 5.1 positions the findings of this paper in the literature.

One of the questions of this study is to find how divers and heterogeneous this elasticity can be with respect to different household characteristics. In order to accurately look at the heterogeneities one needs to include the interactions of all the controls with the house price variable in the regression, however, this has not been the common practice. For instance, to estimate the heterogeneity in the elasticity by income, common practice has been to interact the house price growth with the income variable and interpret the estimated coefficient as the heterogeneity by income. The estimated coefficient is however contaminated by the effect of other variables such as age that has not been taken into account; high income households may be mostly old and may have different elasticities compared to younger households.

[17] As before, column (1) does not include any household controls, column (2) includes household fixed effects and column (3) replaces the fixed effects with controls.

Therefore, a more accurate way of estimating the heterogeneities is to control for the interactions of all of the controls and the endogenous variable in the regression as shown below.

$$\Delta c_{it} = \sum_{p=1}^{P} \nu_p X_p + \sum_{p=1}^{P} \theta_p X_p \Delta hp_{ct} + \epsilon_{it} \tag{6}$$

In the above equation, p is the index for each covariate for a total of P variables, θ is the heterogeneity parameter, X is the control variables and the notation for consumption and house price growth are as before. The number of coefficients to be estimated in this regression is $2P$ plus the fixed effects. With limited number of observations, this regression does not have enough power to estimate all of the coefficients. Moreover, including the interactions linearly may still pose some concerns. For instance the relationship between income and age may not be linear and therefore, the regression needs to include both the first order and the second order interactions of the covariates. In this case, estimating the heterogeneities requires the following functional form.

$$\begin{aligned} \Delta c_{it} = \sum_{p=1}^{24} \nu_p X_p + \sum_{p=1}^{24} \theta_p X_p \Delta hp_{ct} + \\ \sum_{p=1}^{24} \nu_{pp'} \prod_{p'=1}^{24} X_p X_{p'} + \sum_{p=1}^{24} \theta_{pp'} \prod_{p'=1}^{24} X_p X_{p'} \Delta hp_{ct} + \epsilon_{it} \end{aligned} \tag{7}$$

The above regression has $2(p + p^2)$ coefficients not including the fixed effects. For example, in the case of having 24 variables, one needs to estimate 820 coefficients plus the fixed effects which can have very low power with limited observations. This is called the "curse of dimensionality" and is more pronounced with more flexible regressions. There exists some econometrics models that try to overcome this issue such as local linear regressions (LLR). In this paper, I use the GRF to estimate a the elasticity as a function of all the covariates. GRF is a machine learning method that is more flexible and powerful than the existing LLRs and kernel regression methods. Section 4.2 describes the advantage of GRF over LLR. But before that the next section provides more details about the GRF model itself and the estimation procedure.

4 Generalized Random Forest (GRF)

AI started in the 50s with the first papers[18] that introduced artificial intelligence to the world. From then on more complicated and application oriented machine learning methods have been developed in computer science. These models have been rapidly growing in influence and become vital parts of many science fields and also businesses.

[18] Pioneers are Alan Turing, Marvin Minsky and Dean Edmonds

ML algorithms have some advantages over common regression methods[19]. They are highly non-linear models with high prediction power. They do not impose any strict functional form, the relationships are non-parametrically derived from the data and tailored to the data at hand. Most of them perform variable selection internally which mitigates the curse of dimensionality.

A significant advantage of ML to OLS methods is that with machine learning one can estimate individual level coefficients which is not possible through regression. This feature allows researchers to look at the entire distribution of heterogeneities in the estimates as opposed to just one point. For instance, instead of looking at the coefficient of the interaction between a time dummy and the variable of interest that will just provide an average comparison of the before vs after, one can precisely look at the distribution of this effect through time using ML.

Considering the features of ML algorithms, it is surprising how these theories and tools have not been incorporated into econometrics and applied economics as in other fields. There are two main previously existing issues that caused this delay: one, lack of large enough datasets and second, lack of econometric knowledge to draw causal inference using more complicated statistical methods such as ML algorithms.

ML models perform best with large datasets. They are complicated models that may overfit the data and subsequently increase the variance of the estimates. In order to make sure that the trained model is not overfitting, one needs to reserve a part of the data for testing and train the model only on a subset of the data. Therefore, if the data is small (the rule of thumb is bellow a couple thousand observations), the model performance may be low. But nowadays owing to the widely available detail (big) data, lack of data is less of a restriction and consequently more new casual machine learning methods have come to the forefront.

It was not until recently that machine learning has been re-introduced by a series of Susan Athey and coauthor's papers[20] to economists. As mentioned, ML had not been commonly used in economic analysis mainly for the lack of ML methods that allow causal inference. Recently, Athey et al. (2019b) introduced a method called Generalized Random Forest (GRF) used in this paper.

GRF is based on a two step IV methodology. The idea is brilliant and simple, it performs a local maximum likelihood where weights are determined by an "adaptive weighting function" derived from a random forest model. They extend the well-known random forest method[21] (developed by Breiman (2001)). They change the random forest

[19] Methods such as Ordinary Least Squares (OLS) and other mutations of it such as weighted least squares, etc.

[20] Examples of the prominent papers are Athey (2015), Athey (2018),Athey et al. (2019a),Athey et al. (2019b) and Wager and Athey (2018).

[21] A random forest model contains some regression trees built on bootstrapped samples. Besides using the bootstrapped sample, trees split using a random sample of variables and not all. This is a small modification over the usual regression tree or bagging (bootstrap aggregation) that decorrelates the

moment condition which is the mean of the predicted outcome to a more general form that can take any function. The next section will provide the formal definition of the GRF model.

4.1 GRF Model

The GRF model is an extension of the Random Forest (RF) model[22]. Random forest and GRF are very similar in that they both perform recursive partitioning, sub-sampling, and random split selection, the difference is going to be in the last step where random forest just takes a simple average of the predictions for each observations but GRF perform a weighted kernel estimation.

RF models estimate the expected value of Y_i; $\mu(x) = E[Y_i|X_i = x]$ for a data generating process of the form $(X_i, Y_i) \in (\chi \times R)$. GRF extends RF's moment condition to the following:

$$E[\psi_{\theta(x),\nu(x)}(O_i)|X_i = x] = 0 \quad \text{for all } x \in \chi \text{ and } i = 1 \dots n \tag{8}$$

Where O_i is the set of observables, $O_i = \{\Delta c_i, \Delta hp_i, Z_i\}$, which consists of the dependent variable, the endogenous variable and the instrument. Respectively, Δc_i is the growth in consumption expenditures, Δhp_i is the growth in HPI, and Z_i is the instrument. X_i is the set of auxiliary covariates, and $\nu(x)$ is an optional nuisance parameter. The parameter of interest is going to be $\theta(X)$ which is the house price elasticity of consumption in this context. The estimated elasticity, $\hat{\theta(x)}$, is therefore a function of all the covariates.

The structural model that I estimate is the following:

$$\Delta c_{it} = \nu(X_i) + \Delta hp_{ct}\theta(X_i) \tag{9}$$

Where $\theta(X_i)$ is the causal effect of house price changes on consumption expenditures and it varies by X_i. Δhp_i is the endogenous house price which I instrument for with Z_i. As before, the nuisance parameter, $\nu(X_i)$, is the intercept.

In a RF model, the statistician observes the true outcome and trains the model to minimize the error of prediction. For instance, if one wants to predict households' consumption, the objective function can simply minimize the mean squared error of the prediction. However, if the task is to estimate an unobserved parameter such as the house price elasticity of consumption, the true parameter is not observed and therefore, there does not exists an error to minimize. GRF overcomes this problem by deriving the moment functions from the conditional exclusion restriction and the conditional orthogonality of errors assumptions described in what follows.

trees and improves the prediction accuracy and performance of the model. For more description of the regression trees, bagging and random forest models look at the Appendix E.

[22] I mostly follow the notation in Athey et al. (2019b) but change the variable abbreviations used in this paper. Also for more information about RF models, please look at appendix E.

The moment conditions used in the GRF model is coming from the following orthogonality conditions; the exclusion restriction $E[Z_i, \epsilon_i|X_i = x] = 0$ and the conditional orthogonality of errors $E[\epsilon_i|X_i = x] = 0$. Corresponding to each assumption, the moment function is presented below.

$$\psi_{\theta(x),\nu(x)} = \begin{Bmatrix} Z_{ct}(\Delta c_{it} - \Delta hp_{ct}\theta(x) - \nu(x)) \\ \Delta c_{it} - \Delta hp_{ct}\theta(x) - \nu(x) \end{Bmatrix} \tag{10}$$

The objective is to minimize the conditional expectation of the moments by finding the parameters that make the expectations of moments as close as possible to zero. Hence, the parameters are derived from minimizing the weighted moment conditions as the following.

$$(\widehat{\theta}(x), \widehat{\nu}(x)) \in \underset{\theta,\nu}{argmin} \left\{ \left\| \sum_{i=1}^{n} \alpha_i(x)\psi_{\theta,\nu}(O_i) \right\|_2 \right\} \tag{11}$$

As the above expressions presents, the parameters are the argmin of an L2 norm of the weighted moment conditions where $\alpha_i(x)$ is the weighting function derived from a random forest. Therefore, after defining the moments, the goal is to find a weighting function that weights households with similar elasticity ($\theta(x)$) more heavily.

In estimating θ at a specific value of x, one needs to define neighboring observations and give higher weights to closer neighbors and lower weights to further observations. As mentioned before, it is not possible to find the weighting function and estimate the parameter using local linear regressions if neighborhoods are defined based on many dimensions, particularly where neighbors are defined based on the entire vector of covariates. After finding the data-adaptive weights for each observation ($\alpha_i(x)$) using RF, the weighted moment condition is the weighted sum of the moment function ($\psi_{\theta(x),\nu(x)}$);

$$\Sigma_{i=1}^{n}\alpha(x : X_i)\psi_{\theta(x),\nu(x)}(O_i) = 0 \tag{12}$$

The estimated $\theta(\hat{x})$ is a conditional local average treatment estimation of the house price elasticity. The next section depicts the advantages of this model over local linear regressions.

4.2 Econometric Advantages of GRF

What are the technical advantages of GRF over local linear regressions (LLR)? First, one can find neighboring observations based on many dimensions in GRF but in LLR neighbors are defined based on only one running variable (such as distance to the boundary). Second, the forest learns the non-linearity patterns in the data and specializes the weights to the specific data at hand. This produces what Susan Athey calls a *'well-targeted'* weighting function.

What are the empirical benefits of running a GRF? The main benefit is the absence of curse of dimensionality that allows me to estimate multidimensional heterogeneities. With a linear IV regression one can only look at one or two heterogeneities at a time and needs to run different models for different heterogeneities of interest. This can pose the specification search problem and also makes the estimates less comparable with each other since they are outcomes of different separately estimated models. On the other hand, in a GRF model, one can look at all the possible heterogeneities and their joint distributions that are estimated using the original model. All the estimates are coming from the same model and are perfectly comparable and there is less concerns about specification search since the correlations are derived from the data and is not subject to the researcher to choose one or the other.

5 GRF Estimation Results

5.1 Distribution of the Elasticities and Important Determinants

I estimate the following GRF model of house price elasticity of non-durable consumption expenditures for each observed household in the data.

$$\Delta c_{it} = \nu(X_{it}) + \theta(X_{it})\Delta hp_{ct} + \epsilon_{it}$$

Where c_{it} and hp_{ct} are growth of real consumption expenditures and HPI. $\nu(X_{it})$ is the nuisance parameter and $\theta(X_{it})$ is the house price elasticity of consumption as a function of controls. Figure 2 shows the distribution of the estimated elasticities. This distribution has a mean of 0.083 and a standard deviation of 0.035[23], close to the average IV estimate.

Among all the controls included in the model, average age of the household heads, debt to income of the CBSA, the average income of the household heads and household size seem to contribute more to estimating the heterogeneities of this elasticity. Figure 1 shows a visualization of the importance matrix calculated based on the main GRF model[24]. The variables that exclusively define household structures are shown in green and the variables that are more related to the financial status of the household are shown in orange.

Based on the importance matrix illustrated in figure 1, household structure variables (excluding race) are important predictors of the heterogeneities in house price elasticity of consumption expenditures. Among which age, household size and having a child under

[23]The data preparation for the GRF model is explained in Appendix D.

[24]In an importance matrix, variables are sorted based on their contributions in reducing the error of prediction and in the case of the GRF, based on their contribution to the moment function. At each split of the tree, the variable that helps the tree to reach a moment closer to zero is going to be ranked as more important. The final importance matrix is based on all the trees in the forest.

18 years old are the most important. On the financial side, variables such as income and employment status are playing an important role in defining the elasticity.

Being able to look at the distribution of the elasticities in figure 2, I can compare my estimates with the literature's findings. The first interesting observation is that although all of these papers are using very similar methodologies (IV regressions with the supply elasticity instrument), they are finding different elasticities in terms of magnitude. Besides using different consumption data, one discrepancy could be that these papers are looking at different households with different household structures and therefore, finding different magnitudes.

In the next section, I look at the characteristics level heterogeneities and reveal some under-explored patterns. I calculate weighted conditional averages of the elasticity for each household and therefore compare the same household with itself.

5.2 Characteristics Level Elasticities

To explore the dimensions of heterogeneity in the house price elasticity, I calculate a population weighted average of the elasticities of the unique household types in the data. I explain the calculations using a simple example below.

Imagine that one wants to know the heterogeneity of the elasticity in the dimension of having a child under 18 years old. The statistic of interest in this case is the difference that having a child under 18 years old makes for the elasticity ($\widehat{\theta}(X_{child}) - \widehat{\theta}(X_{no\,child})$). An important point to notice here is that the desired estimand is not the average of the elasticity for households with and without a child. Simple averaging will lead to biases due to other household characteristics such as age. Having a child is correlated with age, and averaging the elasticity for households that have a child is analogous to averaging the elasticity for young households with high elasticity. Therefore, with simple averaging the observed difference may be contaminated with the effect of age.

To avoid this issue, I predict the elasticities for all the unique households in the data, first assuming that they all have a child and then repeat the exercise assuming that they all do not have a child under 18. I then weight each unique household by the population weight and calculate the difference. The difference between the weighted average elasticity of the first estimation and the second shows the contribution of this variable to the elasticity. The defined difference for this example will formally be the following.

$$\widehat{\theta}(X_c) = \sum_{h} \gamma_h * \widehat{\theta}(X_c)_h \tag{13}$$

Where h is each unique household type in the population and the difference $\widehat{\theta}(X_{child}) - \widehat{\theta}(X_{no\,child})$ compares the same households with and without a child under 18. With comparing the same household with itself, I am fixing other household characteristics and only look at the change in the desired dimension which does not have the contamination problem described previously.

In the above exercise, since I am comparing the same household with itself, the covariance between $\widehat{\theta}(X_{child})$ and $\widehat{\theta}(X_{no\ child})$ is not zero and hence calculating the variance of the difference is not straight forward[25]. I therefore bootstrap the variance of this difference in three steps. I first train 50 models on bootstrapped data with replacement. Second, I estimate $\widehat{\theta}(X_{child}) - \widehat{\theta}(X_{no\ child})$ as explained above using each of these 50 models and finally calculate the variance of the estimated differences[26].

In the following subsections, I explain the findings based on heterogeneities estimated using the method that I described above. I divide the heterogeneities into household structure heterogeneities and household financial variables heterogeneities.

5.2.1 Household Structure Heterogeneities

I categorize age, having a child under 18, being married and household size as household structure variables. Figure 3, shows the estimated differences between sub-groups of each variable. Among all of the household structure controls, age and presence of a child under 18 years old are the most important determinants of the house price elasticity.

The first part of the graph is comparing young, middle-aged and old household heads. Age is the average age of the household heads. To estimate the elasticity for young, middle-aged and old households, I replace the age variable for all the unique households by 40, 60, and 75 respectively and follow the estimation procedure described above for both the estimated differences and the bootstrapped variances. Young households' elasticities are on average 37%[27] higher than old households by 0.027. . However, there is no significant difference between the middle-aged and older households level of response to house prices.

The second part of figure 3 shows how the average elasticity differs if a household does or does not have a child under 18. The average elasticity for households with and without a child under 18 years old is respectively 0.103 and 0.078. The difference of 0.025 in elasticity means that in the case of a 1% increase in house prices, households with a child under 18 years old, keeping all the other characteristics of the household constant, respond by 35% more than households without a child based on the estimates of table 6.

Marital status is not a significant determinant of house price elasticity of consumption after controlling for all other household characteristics. This finding suggests that the elasticity does not simply increase by adding more people to the household. In other words adding an adult to a family is different from adding a child.

[25] $var(\widehat{\theta}(X_{child}) - \widehat{\theta}(X_{no\ child})) = var(\widehat{\theta}(X_{child})) + var(\widehat{\theta}(X_{no\ child})) - 2cov(\widehat{\theta}(X_{child}), \widehat{\theta}(X_{no\ child}))$

[26] The number of times that I bootstrap is coming from a trade-off between computational cost and precision benefit.

[27] The percentage is calculated based on the level estimates for old households presented in table 6 by dividing 0.027 by 0.072.

To investigate the role of children even further, the last section of figure 3 shows the difference between average elasticities of a single individual and a family of two (first bar) and the difference between a family of two and three (second bar). There is no significant difference in the elasticity of a single individual and a family of two which is consistent with the findings on marital status. However, moving from being single to a family of three, the elasticity increases from 0.076 by 0.014 to 0.09 and is highly significant. Married households with no child are not more sensitive than single individuals but they are more sensitive after the first child by 0.014 and if this child is younger than 18, the difference is even higher.

This is a very interesting finding that has not been directly studied in the house price elasticity literature before. There may be two main reasons for this difference, one is that households with more children (specifically younger children) are more financially constrained and in the event of a positive house price shock, they are going to be more responsive. The other reason may be that it is harder to smooth consumption when there is a child in the family.

An evidence inline with the first hypothesis can be Burrows (2018)'s finding in support of the importance of family composition in home equity withdrawal. They estimate a logit model of the probability of mortgage equity withdrawal and find that households with children are more than twice as likely to withdraw equity than households without children. They find this evidence to hold for both married and single parents. From the collateral channel perspective, if households with younger children are more likely to withdraw equity from their houses, they are more likely to increase consumption as a response to a house price shock which is consistent with the overall message of figure 3.

Older studies such as McClements (1977) emphasized on the importance of the presence and the age of children in families living standards and welfare[28]. Bollinger et al. (2012) use the British Household Panel Survey to show that there exists economies of scale from cohabiting with partners but substantial diseconomies of scale from the first child. They show that addition of the first child increases costs to an extent that the household needs to have almost double the income to keep the same welfare but adding more children increases the costs more mildly.

Prior studies in the house price elasticity literature have not focused on household composition as a determinant of the elasticity due to restrictions in data and empirical approach. However, the heterogeneities in household composition is revealed by the fact that I am estimating the house price elasticity as a function of all the observed household characteristics and also because these estimated heterogeneities are data driven. In the next section, I explain my findings on the heterogeneities in the household financial dimensions.

[28]McClements (1977) is developing an equivalence scale (a measure for cost of living) taking into account the number and ages of children.

5.2.2 Household Financial Structure Heterogeneities

I consider educational attainment, employment status, occupation type of household heads and income as factors that shape households financial structure. Using the same methodology as before (comparing the weighted average of the elasticity of the unique households changing the variable of interest and bootstrapping for the variance of the difference), I study the role of each of these variables in the heterogeneity of the house price elasticity below.

Education levels of household heads is not a significant determinant of the house price elasticity. The first part of figure 4 shows that the average difference between households with different educational attainments is not significant. There is no paper that directly studies the correlation between education levels and house price elasticity, however, Burrows (2018) find a correlation between refinancing and education. They find that the probability of refinancing is higher for higher educated households. Through the collateral channel one can translate their finding to higher house price elasticity of consumption for higher educated households which contrasts my findings. One explanation can be that they are looking at only one dimension of the heterogeneity without considering the interactions between education, income, occupation and other characteristics of households. More educated households are more likely to have high paying jobs and therefore the correlation may be through income rather than education.

The second and third part of figure 4 show the differences for occupation and employment categories. I find no difference between having blue collar or white collar jobs but households in which the head of the household does not work (either a housewife, retired, unable to work or laid off) are less sensitive than households with white collar jobs[29]. A similar pattern is observed for employment. House price elasticity is not different for households with both heads employed rather than one head. However, households with two employed heads are more responsive than households with both heads not working for pay.

Consistently, low and middle income households[30] are not differently responsive to house price shocks but the elasticity for high income households is higher than low income households by 0.01. This is in contrast with the findings of the literature which consists of papers that use a linear IV regression and interact income with the house price variable. Looking at the income heterogeneity without considering all other characteristics that may be correlated with income such as age, education, occupation, etc. one finds a different direction for this effect. It has been commonly accepted that high income is correlated with low elasticity and the explanations provided so far focused on the collateral

[29] I observe the occupation of household heads in 12 levels which I consequently divide into 3 categories; white collar, blue collar and not working. I provide more details about how I categorize the reported occupation levels to blue, white and not working categories in Appendix D.

[30] Low, middle and high income is defined based on the first, second and third quartile of normalized average income and set to $20,000, $50,000 and $80,000.

channel. They argue that high income households are less financially constrained and therefore less responsive to house price shocks. Therefore, what the literature finds on income is basically an average effect of many different characteristics of the household that can affect the elasticity from different channels rather than only through being financially constraint.

Another evidence for this discrepancy that arises from ignoring other dimensions of heterogeneity can be seen in figure 9[31]. Other papers that look at one dimensional heterogeneities usually find that locations with higher income have lower elasticity which is consistent with what this map shows, however, looking at figure 5, the pattern is reverse when the income effect is isolated from other characteristics of the household and location. This finding emphasizes the importance of using a comprehensive model for studying the heterogeneities.

My finding of higher elasticity for higher income levels can also be another evidence for the newly developed literature on “wealthy hand to mouth”. Studies such as Kaplan et al. (2014) and Olafsson and Pagel (2018) show that some households hold little or no liquid wealth despite owning considerable illiquid assets such as housing which suggests high transitory or house price shock responses by these households. The wealthy hand to mouth can be one reason why households in the forth quartile of the income distribution are more sensitive than households in the first quartile in my analysis.

I find no heterogeneity based on the type of residence. As in figure 6, there is no significant difference in the elasticities based on the type of residence holding all other characteristics of the household constant. And no difference between the responsiveness of different races except for Black and African-American and White where the later have a smaller elasticity. There may be some unobservables that are correlated with race that may be responsible for this difference. The observed heterogeneities in the elasticity are only suggestive correlations and are not causal. Even though they are more precise compared to a linear model that is only looking at one or a selected set of dimensions for the heterogeneity.

6 Policy Implications

My results can have two main implications for macro modeling and policy making. First, knowing the heterogeneities in the elasticity can help improving macroeconomic models. I show that having a child is an important determinant of the house price elasticity of consumption. This information is easy to collect and is easy to incorporate into modeling utility functions. Accounting for the heterogeneities can make the models more realistic

[31] This map shows the weighted average of the household level elasticity for 2004 to 2016 at the state level.

and can reveal new patterns[32].

Second, I provide direct policy implications for any policy that affects house prices or access to housing equity. In general, knowing the heterogeneities of the consumption response helps economists and policymakers understand the impacts of policies that directly (issuing more permits) or indirectly (monetary policies that affect interest rates) affect housing markets and house prices.

Being able to estimate the house price elasticity of consumption as a function of household controls allows for the prediction of the elasticity for any location over time. Consequently, one can estimate the consumption expenditure response to house price shocks for each location and time. I provide the average estimated elasticities at the county and state level and follow by some specific policy implications below.

As mentioned before, after estimating the elasticity as a function of household characteristics, $\widehat{\theta}(X)$, one can predict the elasticity for any household in any location and time. To illustrate this feature of my results, I predict the elasticity for households in my original raw survey data and calculate the weighted average for each county. Figure 8 shows the map of the weighted average of the household level house price elasticity of consumption from 2004 to 2016 at the county level.

It is important for policymakers to take both local and regional heterogeneities of the elasticity into account. Figure 8 reveals some interesting patterns related to this point. First, neighboring counties within a state can have very different elasticities (local heterogeneity). This fact should be considered by policymakers when deciding about housing related policies to prevent unintended welfare consequences. Second, looking at weighted average elasticities at the state level (figure 9), there is a great regional heterogeneity as well as local in the average elasticity, west region have lower elasticities compared to the east on average.

There are a range of policies for which knowing the heterogeneities in consumption response matters for policy making. Housing policies that affect access to credit can have heterogeneous welfare consequences due to the heterogeneities in the elasticity. For example, a 2010 law in Singapore increased the minimum necessary occupation period for public housing apartments from 3 to 5 years which meant that households could not access their housing stock for a longer period of time. Agarwal and Qian (2017) use this experiment and show that after the negative credit access shock, households decreased their credit card consumption by 27% in the 6 months after the policy and this reduction was mainly driven by non-durable consumption. Such policies may be effective for objectives such as controlling a booming housing market but may also have significant negative welfare effects for the population that policymakers want to help in

[32] To model intertemporal optimization in consumption growth, Attanasio and Weber (1995) enter having children in different ages as demand shifters in the utility function. They discuss that utility from a certain consumption expenditure is not independent of household composition and failure to consider individual level heterogeneities can lead to bias estimates of the preference parameters.

the first place. For instance, if households that are living in these public housings are young with more children and their consumption is sensitive to their house prices, then blocking access to credit can have substantial welfare effects for this group.

The next section emphasizes on the importance of heterogeneities in the house price elasticity of consumption for calculating accurate aggregations and drawing inference. It also relates the potential bias to the existing differences in locations with high and low volatility in house prices.

7 Potential Aggregation Bias and the Boom-Bust Asymmetry

At the macro level, policymakers are usually interested to know the aggregate effect of a one percentage change in house prices on aggregate consumption and consequently GDP. This is useful in measuring the aggregate effects of recessions and to determine the magnitude of the necessary monetary stimulus. In the following I contrast two estimation methodologies using a single average elasticity measure for all counties (macro aggregation) and using different elasticities for different counties (micro aggregation).

In order to measure the aggregate consumption response using the existing average elasticity measures provided in the literature, one can calculate the weighted average of the interaction of elasticities with the local house price growth. The following is the aggregate consumption growth estimate disregarding the heterogeneities in the local house price elasticity of consumption (ΔC_{Macro}).

$$\Delta C_{Macro} = E[\Delta HP_{local,t}] * \bar{\theta}_{national} \tag{14}$$

Where $\bar{\theta}_{national}$ is the average national elasticity that is estimated using a linear IV regression. The accuracy of this measure falls on the assumption that there is no covariance between the local elasticities and the local growth in house prices. The following equation presents the aggregate consumption growth considering the local heterogeneities in the house price elasticity (ΔC_{Micro}).

$$\begin{aligned} \Delta C_{Micro} =& E[\Delta HP_{local,t} * \theta_{local,t}] \\ =& E[\Delta HP_{local,t}] * E[\theta_{local,t}] + cov(\Delta HP_{local,t}, \theta_{local,t}) \end{aligned} \tag{15}$$

Comparing equations 14 and 15 and assuming that the average of local elasticities equals to the national estimate, the aggregate consumption response derived from the macro methodology is misspecified if $cov(\Delta HP_{local,t}, \theta_{local,t}) \neq 0$. Figure 10 plots the covariance of the house price growth and the estimated elasticities over time. This covariance is not zero and has an interesting pattern; it is negative during the boom years and positive during the bust years (recession period). Figure 11 shows the estimated ΔC_{Macro} and ΔC_{Micro} over time. The macro aggregation overestimates in boom years and underestimates in busts.

The reason for overestimation during booms and underestimation during busts is that the $cov(\Delta HP_{local,t}, \theta_{local,t})$ is negative for busts and positive for booms. Figures 12, 13, and 14 plot the weighted average elasticities for each county against the county house price growth for years 2005 to 2016. The correlation between the county elasticities and county house price growths is negative for boom years and positive for bust years consistent with the findings on covariance. The elasticity falls with the absolute value of the house price changes. In boom years, higher house price growth corresponds to lower elasticity and in bust years, locations with higher drops in house prices have lowest elasticities. This result is quite surprising since one expects that locations that have been largely affected by the recession to show higher responsiveness in terms of consumption.

Dividing the counties into counties with above and below average variance in house prices (counties shown in triangle and round shapes respectively in figures 12, 13, and 14) reveals an interesting pattern. Counties with high volatile housing markets have low elasticities and bring about the observed negative correlation between the house price growth and elasticities in booms and and the positive correlation in busts. This finding is inline with the county and state level maps of elasticities shown in figures 8 and 9. High house price locations such as the west coast and Florida have low elasticities.

Disregarding the systematic differences and using the same elasticity to calculate the aggregate consumption response introduces biases. High volatile housing markets with large house price changes have low elasticities and using the average elasticity for these locations leads to overestimation in booms (large upticks but low elasticity) and underestimation in busts (large upticks but low elasticity).

Is there asymmetry in the house price elasticity of consumption in terms of booms and busts? The results from the previous part suggests that comparing counties with different house price changes and consequently different house price volatility levels is not the right way of measuring asymmetry. High house price variance locations have lower levels of elasticities in both booms and busts which leads to finding no asymmetry. However, looking at individuals within counties reveals the asymmetric patterns. On average households consumption responses to house prices is larger during bust years compared to booms. Table 8 presents the OLS results of the following regression.

$$\hat{\theta}_{it} = \beta_0 + \beta_1 X_{it} * Bust + \beta_2 X_{it} + \beta_3 Bust + \gamma_c + \epsilon_{it} \tag{16}$$

Where $\hat{\theta}_{it}$ are the estimated household level house price elasticities using GRF, X_{it} is either a dummy for household being young, having a child under 18, and household size. *Bust* is a dummy for years from 2007 to 2012 and γ_c is a county fixed effect.

The first column of this table regresses the household level elasticities on the bust dummy. The coefficient is positive and significant and suggests that on average elasticities are higher by 0.011 (about 13.25% of the average elasticity of 0.083). Based on the GRF results of section 5.2, I consider being young, having a child under 18 and having a large family size as variables that are correlated with being financially constrained and having preferences favoring current consumption. Younger households respond more than older households in general and during busts compared to booms. The same pattern is observed in columns 3 and 4 of table 8 for households with a child and larger households.

8 Conclusion

A rich literature in macroeconomics and household finance has shown evidence for positive marginal propensity to consumption (MPC) out of wealth and house price elasticity of consumption. The collateral and wealth channels are the two main mechanisms by which wealth or house prices can causally change household consumption. Despite the consensus on the existence of this relationship, the literature does not provide a full picture of the heterogeneities in the house price elasticity of consumption and does not emphasize on the importance of regional differences in this elasticity.

Economists and policymakers use the house price elasticity estimates to predict the effect of house price changes on consumption and consequently GDP. But, as this paper shows, average elasticity is not a good benchmark for evaluating local and individual policy impacts. There exists substantial regional and household level variations in the the elasticity which policymakers should consider to evaluate the consequences of their policies for different regions and households. I find that locations with high volatile housing markets are less elastic and disregarding this fact leads to overestimation of the aggregate consumption response during booms and underestimation during busts.

In this paper, I estimate the house price elasticity of consumption as a non-parametric function of the household characteristics, location and time. I find that on average, for a 1% increase in real HPI, households increase their real consumption expenditure by 0.083%. This average derived from a generalized random forest (GRF) model is reassuringly close to the linear IV estimate of 0.09. However, there exists considerable disparities both at the county and household level.

Elasticities are larger for younger households, with a child, and of larger size. Having a child, conditional on all other characteristics, increases the elasticity by 0.025. Employment, occupation and income define the difference between the two tails of the distribution. The elasticity is higher during bust years (2007 to 2012) and larger for more constrained and sensitive households. The heterogeneities and the distribution of the elasticity with respect to fundamental characteristics can have important implications for future research.

References

Agarwal, S. and Qian, W. (2017). Access to home equity and consumption: Evidence from a policy experiment. *Review of Economics and Statistics*, 99(1):40–52.

Aladangady, A. (2017). Housing wealth and consumption: Evidence from geographically-linked microdata. *American Economic Review*, no. 11:3415–46.

Aruoba, S. B., Elul, R., and Kalemli-Ozcan, S. (2019). How big is the wealth effect? decomposing the response of consumption to house prices.

Athey, S. (2015). Machine learning and causal inference for policy evaluation. In *Proceedings of the 21th ACM SIGKDD international conference on knowledge discovery and data mining*, pages 5–6. ACM.

Athey, S. (2018). The impact of machine learning on economics. In *The Economics of Artificial Intelligence: An Agenda*. University of Chicago Press.

Athey, S., Bayati, M., Imbens, G., and Qu, Z. (2019a). Ensemble methods for causal effects in panel data settings. Technical report, National Bureau of Economic Research.

Athey, S., Tibshirani, J., Wager, S., et al. (2019b). Generalized random forests. *The Annals of Statistics*, 47(2):1148–1178.

Attanasio, O. P. and Weber, G. (1995). Is consumption growth consistent with intertemporal optimization? evidence from the consumer expenditure survey. *Journal of political Economy*, 103(6):1121–1157.

Beraja, M., Fuster, A., Hurst, E., and Vavra, J. (2018). Regional heterogeneity and the refinancing channel of monetary policy. *The Quarterly Journal of Economics*, 134(1):109–183.

Berger, D., Guerrieri, V., Lorenzoni, G., and Vavra, J. (2017). House prices and consumer spending. *The Review of Economic Studies*, 85(3):1502–1542.

Bollinger, C., Nicoletti, C., and Pudney, S. (2012). Two can live as cheaply as one... but threeâsa crowd.

Breiman, L. (2001). Random forests. *Machine learning*, 45(1):5–32.

Burauel, P. and Schroeder, C. (2019). The german minimum wage and wage growth: Heterogeneous treatment effects using causal forests. *Available at SSRN 3415479*.

Burrows, V. (2018). The impact of house prices on consumption in the uk: A new perspective. *Economica*, 85(337):92–123.

Campbell, J. Y. and Cocco, J. F. (2007). How do house prices affect consumption? evidence from micro data. *Journal of Monetary Economics*, 54, no. 3:591–621.

Carter, M. R., Tjernström, E., and Toledo, P. (2019). Heterogeneous impact dynamics of a rural business development program in nicaragua. *Journal of Development Economics*, 138:77–98.

Cloyne, J., Huber, K., Ilzetzki, E., and Kleven, H. (2019). The effect of house prices on household borrowing: A new approach. *American Economic Review*, 109(6):2104–36.

Cui, Y. and Davis, A. M. (2019). Tax-induced inequalities in the sharing economy. *Available at SSRN 3376992*.

Daoud, A. and Johansson, F. (2019). Estimating treatment heterogeneity of international monetary fund programs on child poverty with generalized random forest.

Davidoff, T. et al. (2016). Supply constraints are not valid instrumental variables for home prices because they are correlated with many demand factors. *Critical Finance Review*, 5(2):177–206.

Davis, J. and Heller, S. B. (2017). Using causal forests to predict treatment heterogeneity: An application to summer jobs. *American Economic Review*, 107(5):546–50.

Gelain, P., Lansing, K. J., and Natvik, G. J. (2015). Explaining the boom-bust cycle in the u.s. housing market: A reverse-engineering approach.

Graham, J. (2018). Consumption responses to house price heterogeneity. *Working Paper*.

Guren, A. M., McKay, A., Nakamura, E., and Steinsson, J. (2018). Housing wealth effects: The long view. Technical report, National Bureau of Economic Research.

Gyourko, J., Saiz, A., and Summers, A. (2008). A new measure of the local regulatory environment for housing markets: The wharton residential land use regulatory index. *Urban Studies*, 45, no. 3:693–729.

Hastie, T., Tibshirani, R., and Friedman, J. (2013). *The Elements of Statistical Learning: Data Mining, Inference, and Prediction*. Springer.

Hoffman, I. and Mast, E. (2019). Heterogeneity in the effect of federal spending on local crime: Evidence from causal forests. *Available at SSRN 3059137*.

Howard, G. and Liebersohn, C. (2018). The geography channel of house price appreciation.

Kaplan, G., Violante, G. L., and Weidner, J. (2014). The wealthy hand-to-mouth. Technical report, National Bureau of Economic Research.

McClements, L. D. (1977). Equivalence scales for children. *Journal of Public Economics*, 8(2):191–210.

Mian, A., Rao, K., and Sufi, A. (2013). Household balance sheets, consumption, and the economic slump. *The Quarterly Journal of Economics*, 128(4):1687–1726.

Mian, A. and Sufi, A. (2011). House prices, home equity-based borrowing, and the us house- hold leverage crisis. *American Economic Review*, 101 no. 5:2132â56.

Mian, A. and Sufi, A. (2014). House price gains and us household spending from 2002 to 2006. *National Bureau of Economic Research*, No. w20152.

Mian, A. and Sufi, A. (2016). Who bears the cost of recessions? the role of house prices and household debt. *Handbook of Macroeconomics*, 2:255–296.

Olafsson, A. and Pagel, M. (2018). The liquid hand-to-mouth: Evidence from personal finance management software. *The Review of Financial Studies*, 31(11):4398–4446.

Palmer, C. (2015). Why did so many subprime borrowers default during the crisis: Loose credit or plummeting prices? *Available at SSRN 2665762*.

Pope, D. G. and Pope, J. C. (2015). When walmart comes to town: Always low housing prices? always? *Journal of Urban Economics*, 87:1–13.

Saiz, A. (2010). The geographic determinants of housing supply. *Quarterly Journal of Economics*, 125(3):1253â1296.

Steven Manson, Jonathan Schroeder, David Van Riper, and Steven Ruggles (2019). Ipums national historical geographic information system: Version 14.0. Minneapolis, MN: IPUMS, http://doi.org/10.18128/D050.V14.0.

Stroebel, J. and Vavra, J. (2016). House prices, local demand, and retail prices. *National Bureau of Economic Research*, No. w20710.

Sullivan, D. (2014). From food desert to food mirage: Race, social class, and food shopping in a gentrifying neighborhood. *Advances in Applied Sociology*.

U.S. Bureau of Labor Statistics (2018). Consumer price index for all urban consumers: All items (cpiaucsl). retrieved from FRED, Federal Reserve Bank of St. Louis, https://fred.stlouisfed.org/series/CPIAUCSL.

Wager, S. and Athey, S. (2018). Estimation and inference of heterogeneous treatment effects using random forests. *Journal of the American Statistical Association*, 113(523):1228–1242.

Wong, A. (2019). Refinancing and the transmission of monetary policy to consumption.

Appendices

A Tables

Table 1. Demographics of Panelists
The following table shows the demographics of panelists in the final data. Panelists information are provided in the Kilts-Nielsen Consumer Panel Dataset (KNCP) and the description of the variables can be found in Appendix D. The first column shows the dummies, the second column describes the levels of the dummy variables and the last two columns show the frequency and percentage of each category from the entire data.

Variable	Level	Frequency	Percentage
Child Under 18	Yes	100,286	23%
	No	338,180	77%
Married	Yes	273,258	62%
	No	165,208	38%
Type of Residence	Single Family House	358,040	82%
	Two	15,151	3%
	Three+	45,795	10%
	Mobile	19,480	4%
Type of Occupation	White Collar	187,779	43%
	Blue Collar	131,197	30%
	Not Working	119,490	27%
Education	High School	86,964	20%
	Undergraduate	294,546	67%
	Graduate	56,956	13%

Table 2. Summary Statistics
Expenditures is the sum of household level consumption expenditures. HPI is the annual house price index from FHFA. Household income is average of household heads' income. The income itself is calculated as an average of the median of the reported income bins. Age is the average age of the heads of the household and size is the total number of people living in the household. Expenditures, HPI and average income is normalized by CPI of 2004.

	Mean	St. Dev.	Min	Pctl(25)	Pctl(75)	Max
Expenditures	\$5,932	\$2,531	\$983	\$3,804	\$8,027	\$11,608
Expenditures Growth	−0.05	0.23	−2.25	−0.17	0.07	2.19
HPI	372	172	69	266	435	1,434
HPI Growth	−0.01	0.07	−0.56	−0.04	0.03	0.32
Household Income	\$50,903	\$30,425	\$1,968	\$27,296	\$69,939	\$187,448
Age	56	12	34	47	65	78
Size	2	1	1	1	3	9
No. of Households in CBSA	318	460	34	90	328	3,190

Table 3. Counties with Highest HPI Growth for Each Year in the US
The FHFA house price index dataset has data for 382 counties. The growth rate is calculated as change in logarithm of HPI. The 2004 CPI is used to normalize the house price index.

	Year	$\Delta log(HPI)$	HPI	County	State
1	2000	27%	281	NANTUCKET	MA
2	2001	17%	118	GRANT	OR
3	2002	19%	158	NEW YORK	NY
4	2003	12%	292	ST. LUCIE	FL
5	2004	24%	516	CLARK	NV
6	2005	32%	204	STOREY	NV
7	2006	40%	143	YOAKUM	TX
8	2007	20%	123	LOVE	OK
9	2008	16%	131	DICKEY	ND
10	2009	23%	99	FRONTIER	NE
11	2010	16%	129	QUAY	NM
12	2011	18%	140	CAVALIER	ND
13	2012	30%	118	FALLON	MT
14	2013	22%	163	MCHENRY	ND
15	2014	36%	121	KITTSON	MN
16	2015	26%	94	PUTNAM	MO
17	2016	33%	109	JEFF DAVIS	GA
18	2017	39%	116	HOLT	MO

Table 4. First Stage Regressions for the Local House Price Sensitivity IV

$$\Delta hp_{ct} = \beta_1' \hat{\gamma}_{CBSA} \Delta hp_{U.S,t} + \beta_2' X_{it} + \alpha_t + \alpha_{ts} + \alpha_{cbsa} + e_{ict}$$

Where α_t, $alpha_{ts}$ and α_{cbsa} are respectively year, state-year and CBSA fixed effects. Household controls (X_{it}) consists of type of residence, time varying household income, average age, age squared, education, employment and occupation of household heads, size, marital status, and race.

	Dependent variable:	
	$\Delta log(HPI)$	
	(1)	(2)
$\hat{\gamma} * \Delta log(National\ HPI)$	0.010***	0.010***
	(0.001)	(0.001)
Instrument F-Statistics	77.79	85.09
Household Controls	N	Y
CBSA FE	N	Y
Year FE	Y	Y
Year*State FE	Y	Y
Cluster at State Level	Y	Y
Observations	335,508	335,508
Adjusted R^2	0.917	0.924
Note:	*p<0.1; **p<0.05; ***p<0.01	

Table 5. OLS and IV Estimates of

$$\Delta c_{it} = \beta_1 \Delta hp_{ct} + \beta_2 X_{it} + \alpha_t + \alpha_{ts} + \alpha_{cbsa} + \epsilon_{it}$$

with the following first stage

$$\Delta hp_{ct} = \beta_1' \hat{\gamma}_{CBSA} \Delta hp_{U.S,t} + \beta_2' X_{it} + \alpha_t + \alpha_{ts} + \alpha_{cbsa} + \zeta_{ct}$$

For variable descriptions look at 4.

	Dependent variable:			
	$\Delta log(Exp)$			
	OLS	OLS	IV	IV
$\Delta log(HPI)$	0.057***	0.050**	0.119**	0.090*
	(0.020)	(0.025)	(0.044)	(0.048)
Household Controls	N	Y	N	Y
CBSA FE	N	Y	N	Y
Year FE	Y	Y	Y	Y
Year*State FE	Y	Y	Y	Y
Cluster at State Level	Y	Y	Y	Y
Observations	335,508	335,508	335,508	335,508
Adjusted R^2	0.007	0.013	0.007	0.013
Note:		*p<0.1; **p<0.05; ***p<0.01		

Table 6. Population Weighted Average Estimates of Household Structure Variables

This table shows the weighted average estimates of the elasticity by household structure variables. Child is a dummy for having a child under 18 years old, age is the average age of the heads of the household, size is the number of individuals in the household, married is a dummy for marital status and race is self explanatory. For more details on variables look at Appendix D. The point estimates and standard errors for the differences are estimated using the methodology described in subsection 5.2. The first columns contain the level estimates and the later columns provide the differences between the levels of elasticities of each group and the selected base group. Standard errors are provided in parenthesis.

	With Child	No Child	With Child - No Child				
Child	0.103***	0.078***	0.025***				
	(0.014)	(0.014)	(0.007)				
	Young	Middle Aged	Old	Young - Old	Middle - Old		
Age	0.099***	0.077***	0.072***	0.027***	0.005		
	(0.015)	(0.014)	(0.014)	(0.007)	(0.003)		
	One	Two	Three	Two - One	Three - One		
Household Size	0.076***	0.081***	0.090***	0.005	0.014***		
	(0.014)	(0.014)	(0.014)	(0.003)	(0.004)		
	Married	Single	Married - Single				
Married	0.086***	0.078***	0.008				
	(0.013)	(0.015)	(0.005)				
	BlackAfrican American	Race	Asian	WhiteCaucasian	B - W	A - W	O - W
Race	0.101***	0.088***	0.090***	0.081***	0.020**	0.007	0.009
	(0.017)	(0.015)	(0.014)	(0.015)	(0.009)	(0.006)	(0.005)

Note: *p<0.1; **p<0.05; ***p<0.01

Table 7. Population Weighted Average Estimates of Household Financial Variables

This table shows the weighted average estimates of the elasticity by household financial variables. Income, education, occupation and employment is the average based on the reported values for the heads of the household. Debt to income the public version of county-level debt to income ratio using the Federal Reserve Bank of New York and IRS data for income taken from Mian et al. (2013) and available here. Type of residence is type of household's dwelling. For more details on variables look at Appendix D. The point estimates and standard errors for the differences are estimated using the methodology described in subsection 5.2. The first columns contain the level estimates and the later columns provide the differences between the levels of elasticities of each group and the selected base group. Standard errors are provided in parenthesis.

	Low	Medium	High	Low - High	Medium - High		
Income	0.078***	0.083***	0.087***	−0.010**	−0.004		
	(0.014)	(0.014)	(0.014)	(0.005)	(0.003)		
	High School	Undergraduate	Graduate	High School - Graduate	Undergraduate - Graduate		
Education	0.089***	0.079***	0.091***	−0.002	−0.012		
	(0.016)	(0.014)	(0.015)	(0.008)	(0.009)		
	Blue Collar	White Collar	Not Working	Blue - White	Not Working - White		
Occupation	0.089***	0.087***	0.068***	0.002	−0.019***		
	(0.015)	(0.014)	(0.014)	(0.007)	(0.007)		
	Low	Medium	High	Low - High	Medium - High		
CBSA Debt to Income	0.088***	0.086***	0.077***	0.011**	0.008		
	(0.014)	(0.014)	(0.014)	(0.005)	(0.004)		
	Both Heads Employed	One Head Employed	Not Employed for Pay	One - Both	Non - Both		
Employment	0.089***	0.081***	0.072***	−0.008	−0.017*		
	(0.015)	(0.014)	(0.014)	(0.007)	(0.008)		
	One Family House	Two	Three	Mobile	Two - One	Three - One	Mobile - One
Type of Residence	0.084***	0.079***	0.076***	0.076***	−0.005	−0.008	−0.008
	(0.014)	(0.015)	(0.016)	(0.018)	(0.005)	(0.008)	(0.009)

Note: *p<0.1; **p<0.05; ***p<0.01

Table 8. Within County Variations in the House Price Elasticity by Household Characteristics
In this table, each column corresponds to a DID analysis with the household level elasticities being the dependent variable as in the following equation.

$$\hat{\theta}_{it} = \beta_0 + \beta_1 X_{it} * Bust + \beta_2 X_{it} + \beta_3 Bust + \gamma_c + \epsilon_{it}$$

The "Bust" dummy is one for years from 2007 to 2012 and zero for 2005 to 2006 and 2013 to 2016. All regressions include county fixed effects and are clustered at the county level. Household characteristics' definitions can be found in Appendix D.

	Dependent variable:			
	House Price Elasticity of Non-Durable Consumption			
	(1)	(2)	(3)	(4)
Young		0.029***		
		(0.001)		
Child Under 18			0.035***	
			(0.001)	
Household Size				0.011***
				(0.0003)
Bust	0.011***	0.010***	0.010***	0.011***
	(0.0004)	(0.0004)	(0.0004)	(0.0004)
Young:Bust		0.001**		
		(0.0003)		
Child Under 18*Bust			0.002***	
			(0.0003)	
Household Size*Bust				0.001***
				(0.0001)
County FE	Y	Y	Y	Y
Cluster at County Level	Y	Y	Y	Y
Observations	334,779	334,779	334,779	334,779
R^2	0.430	0.815	0.887	0.878
Adjusted R^2	0.429	0.813	0.886	0.877
Note:			*p<0.1; **p<0.05; ***p<0.01	

B Figures

Figure 1. The Importance Matrix Calculated for the GRF Model
This figure illustrates the variables in terms of their importance for estimating the elasticities. Variables are ranked in order of importance from top to bottom and are divided into household structure and financial structure. For description of the variables check Appendix D.

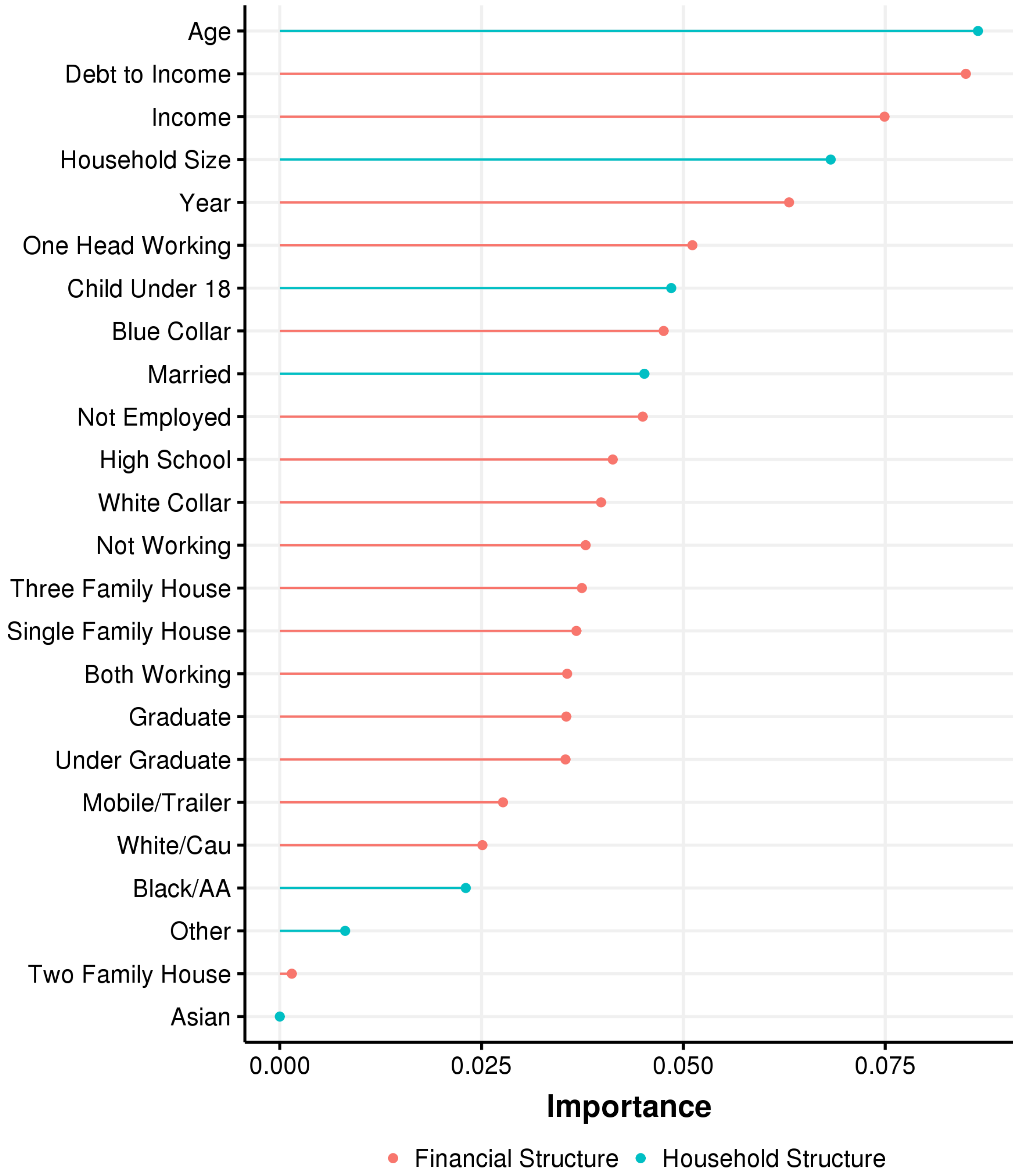


Figure 2. The Distribution of GRF Estimates of Average House Price Elasticities of Consumption Expenditures
This plot shows the distribution of the household level estimations of the house price elasticity of consumption during 2005 to 2016. Elasticities range from 0.01 to 0.21. The standard deviation of this distribution is 0.035. The vertical axis shows the interaction of the density and the width of each bin and the horizontal axis is the GRF's house price elasticity of distribution.

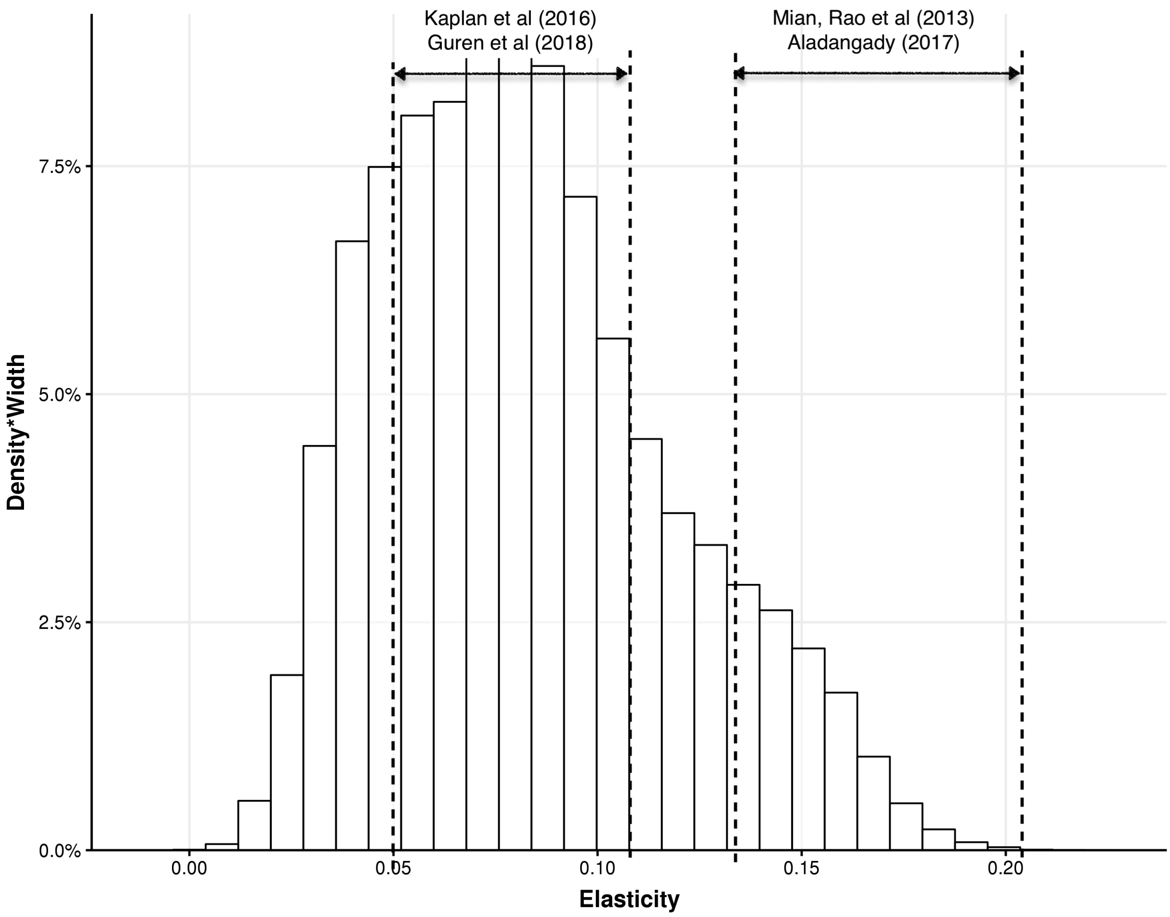


Figure 3. Heterogeneities by Household Structure
This plot consists of four main parts; age, child, married and size. In each part, the two groups mentioned underneath the title are compared to each other. For example, the first bar is showing the difference between the elasticity of being young vs old. Young, old and middle-aged denote the average age of the heads of the household and are defined to be 40, 60, and 75 respectively. Child is a dummy for existing a child under 18 years old in the family. Married is a dummy for marital status and size is the size of the household. 95% confidence intervals are shown.

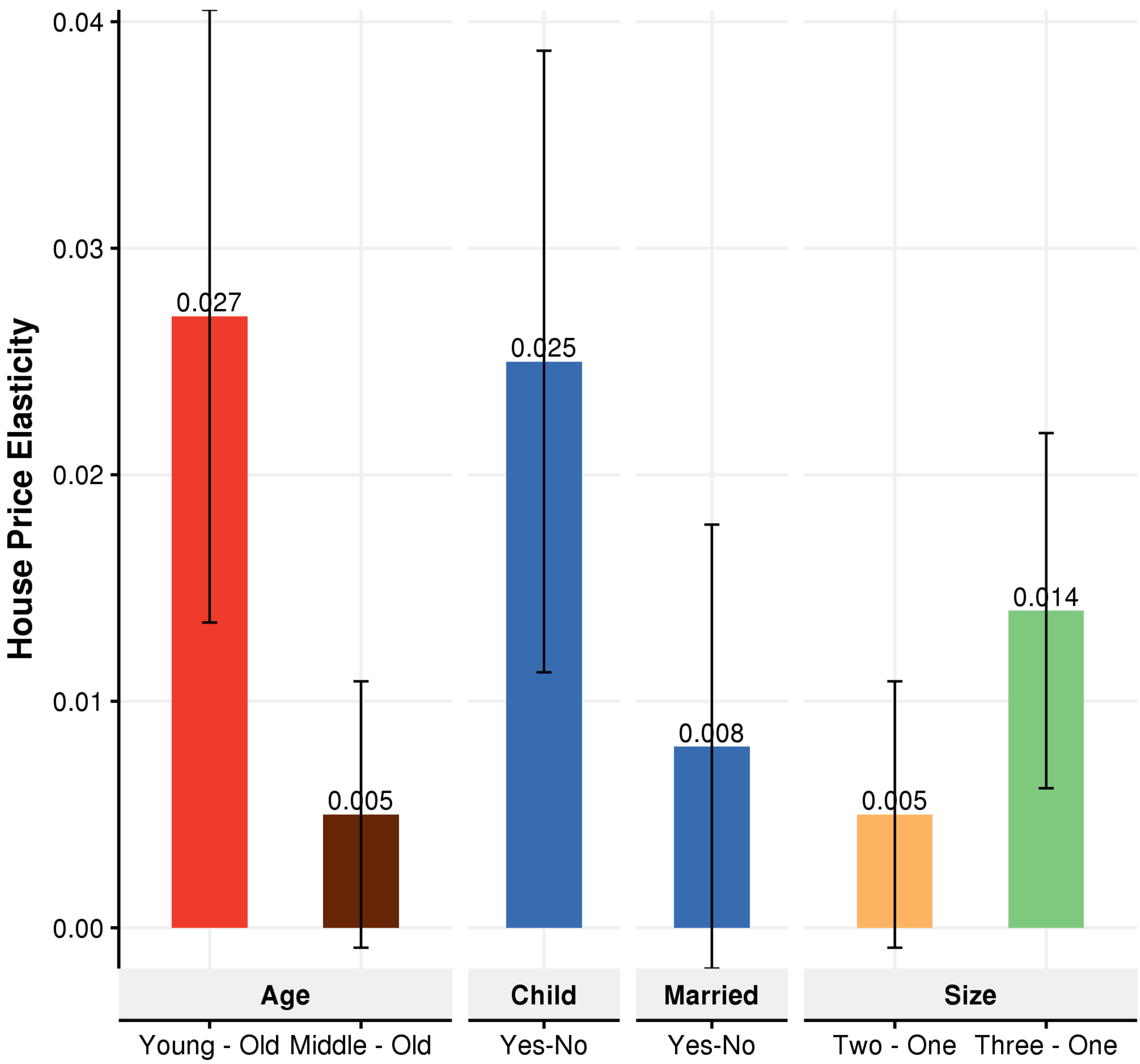


Figure 4. Heterogeneities by Household Financial Structure
This plot has three parts, education, occupation and employment. These variables are defined based on the average characteristics of the heads of the household. Each part of the plot has two subplots which shows the magnitude and the 95% confidence interval of the difference between the two groups denoted under each category. For more details on the variable definitions, check Appendix D.

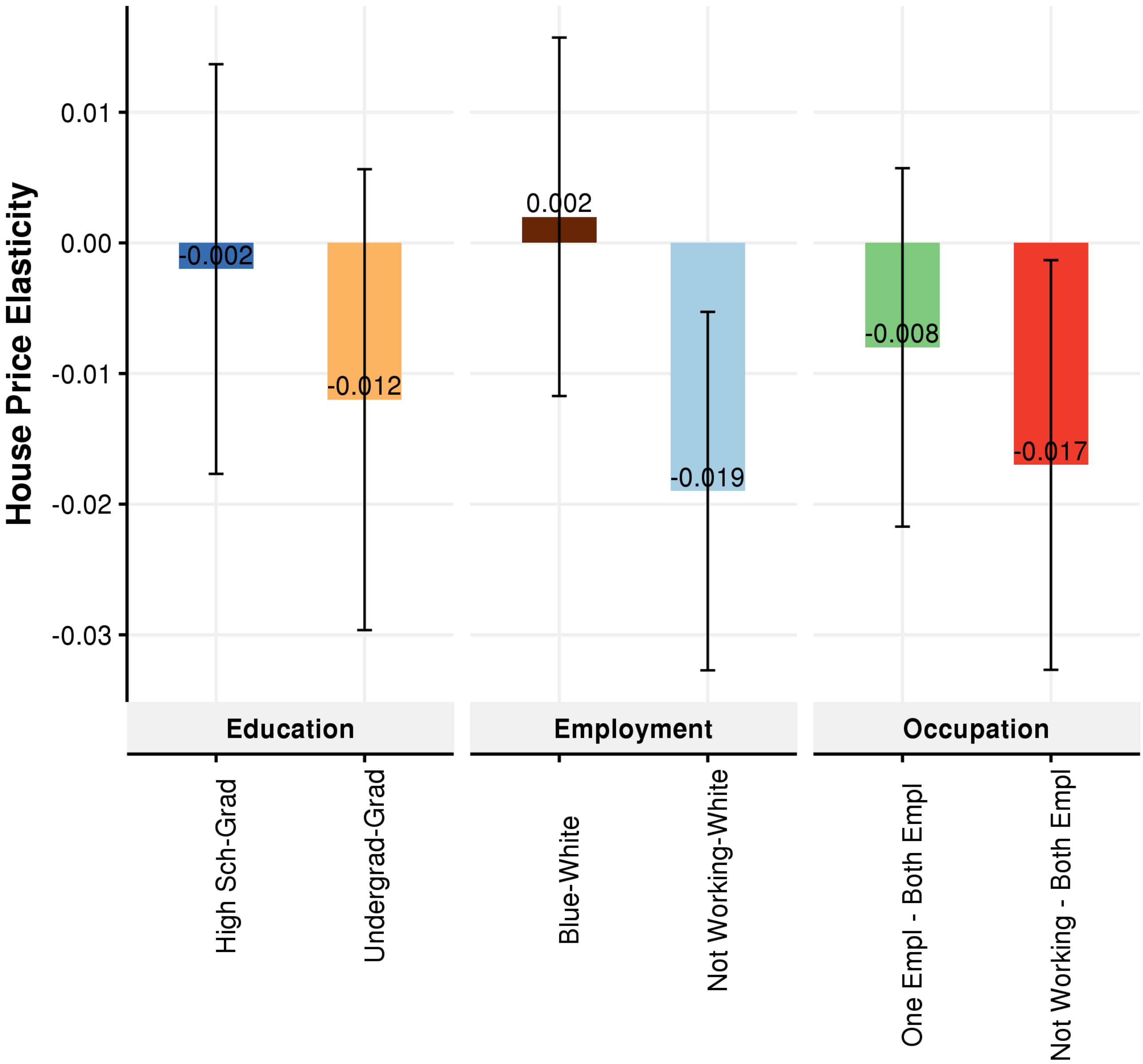


Figure 5. Heterogeneities by Income
Income is defined as the average income of the heads of the household normalized by CPI of 2004. The low, medium and high income levels are set at \$20,000, \$50,000 and \$80,000. Bars show the difference between the elasticity of low vs high and middle vs high income levels. 95% confidence intervals are shown.

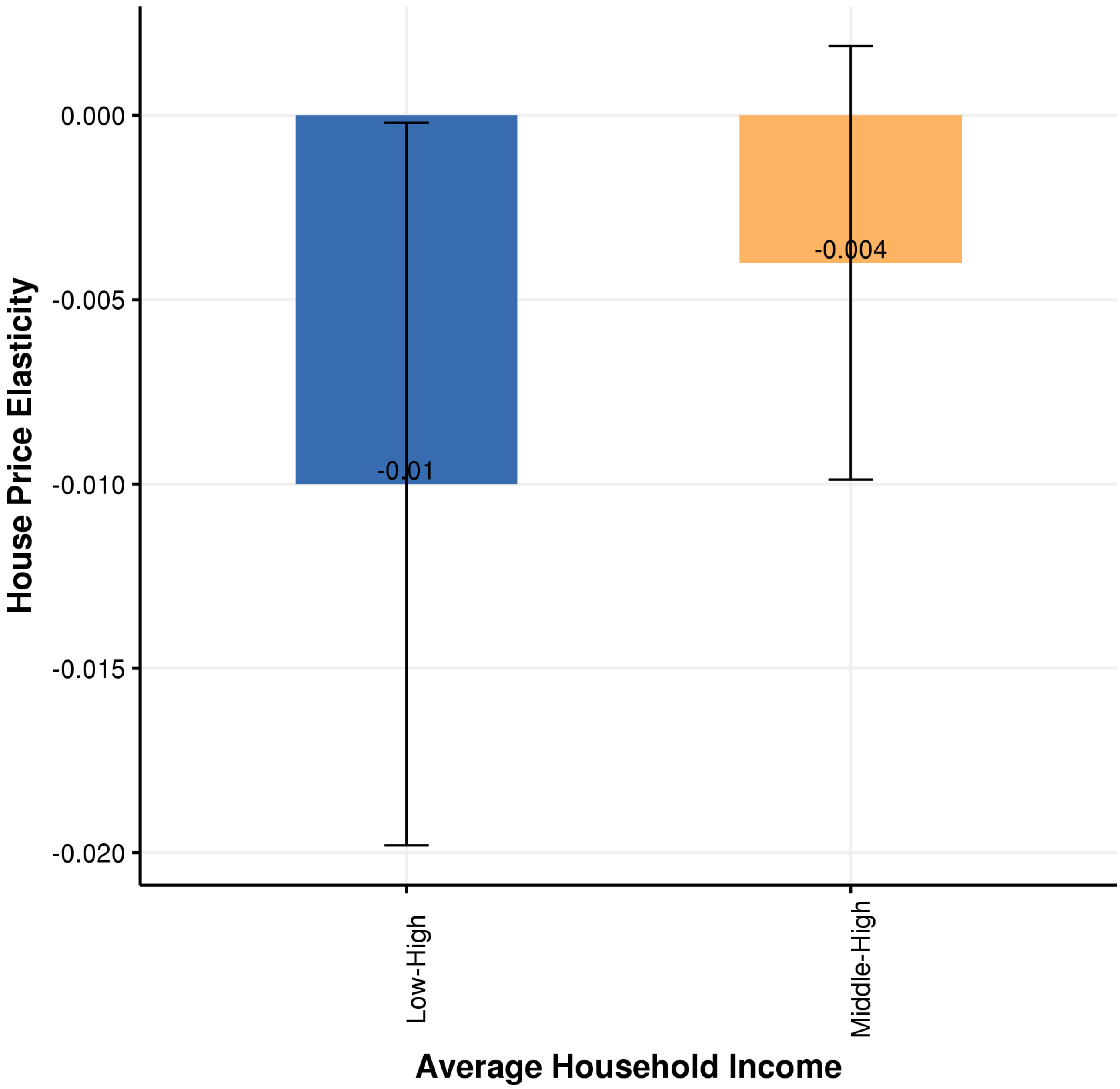


Figure 6. Heterogeneities by Type of Residence
Type of residence is summarized in four categories of single family house, two family house, three plus and mobile and trailer. All the bars are comparing the elasticity of different residence types with a single family house. For more information about variable definitions check Appendix D.

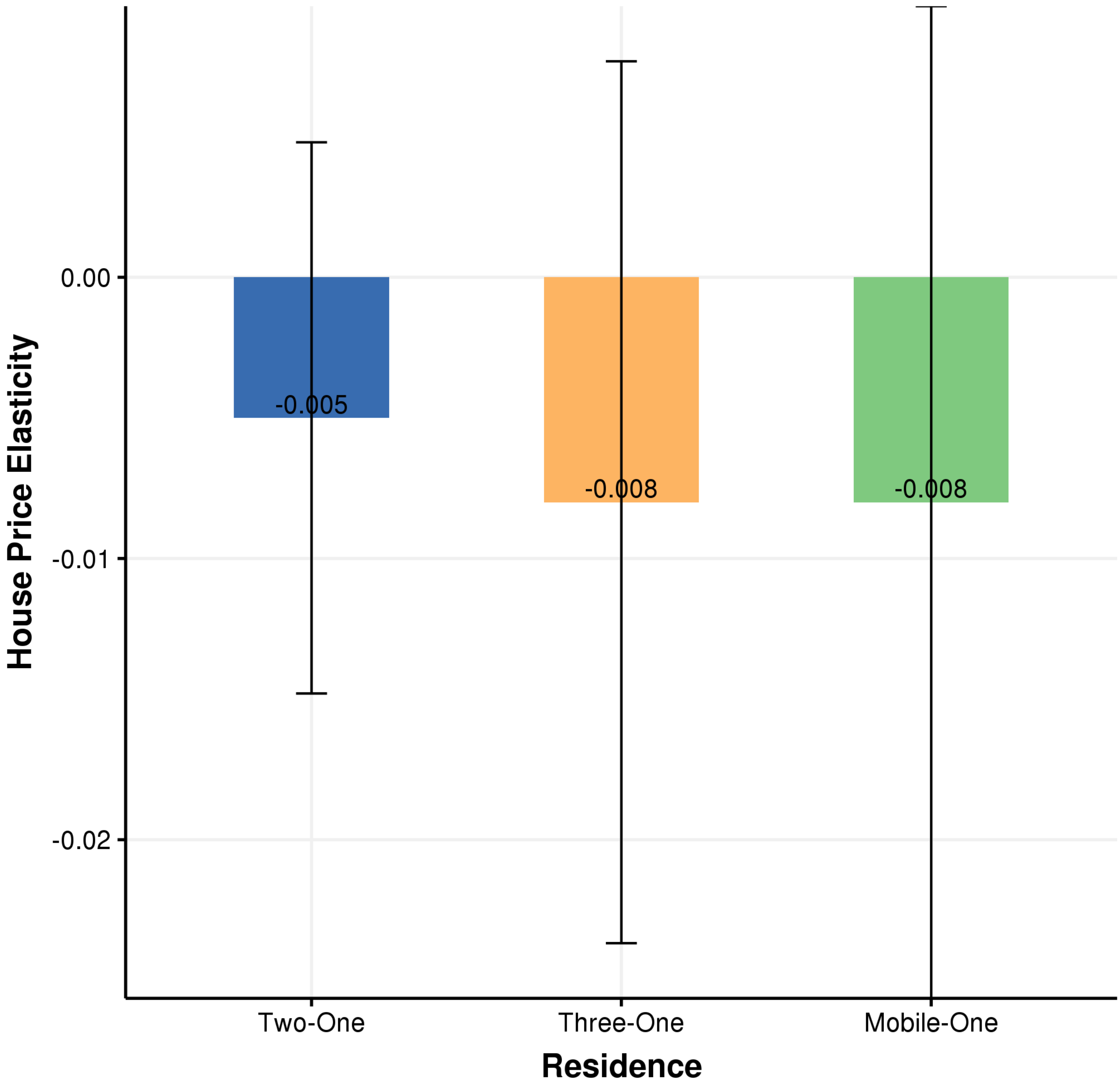


Figure 7. Heterogeneities by Race
Race is reported in four categories Black/African American, White/Caucasian, Asian and other. All the bars are comparing the elasticity of different races compared to White. 95% confidence intervals are shown.

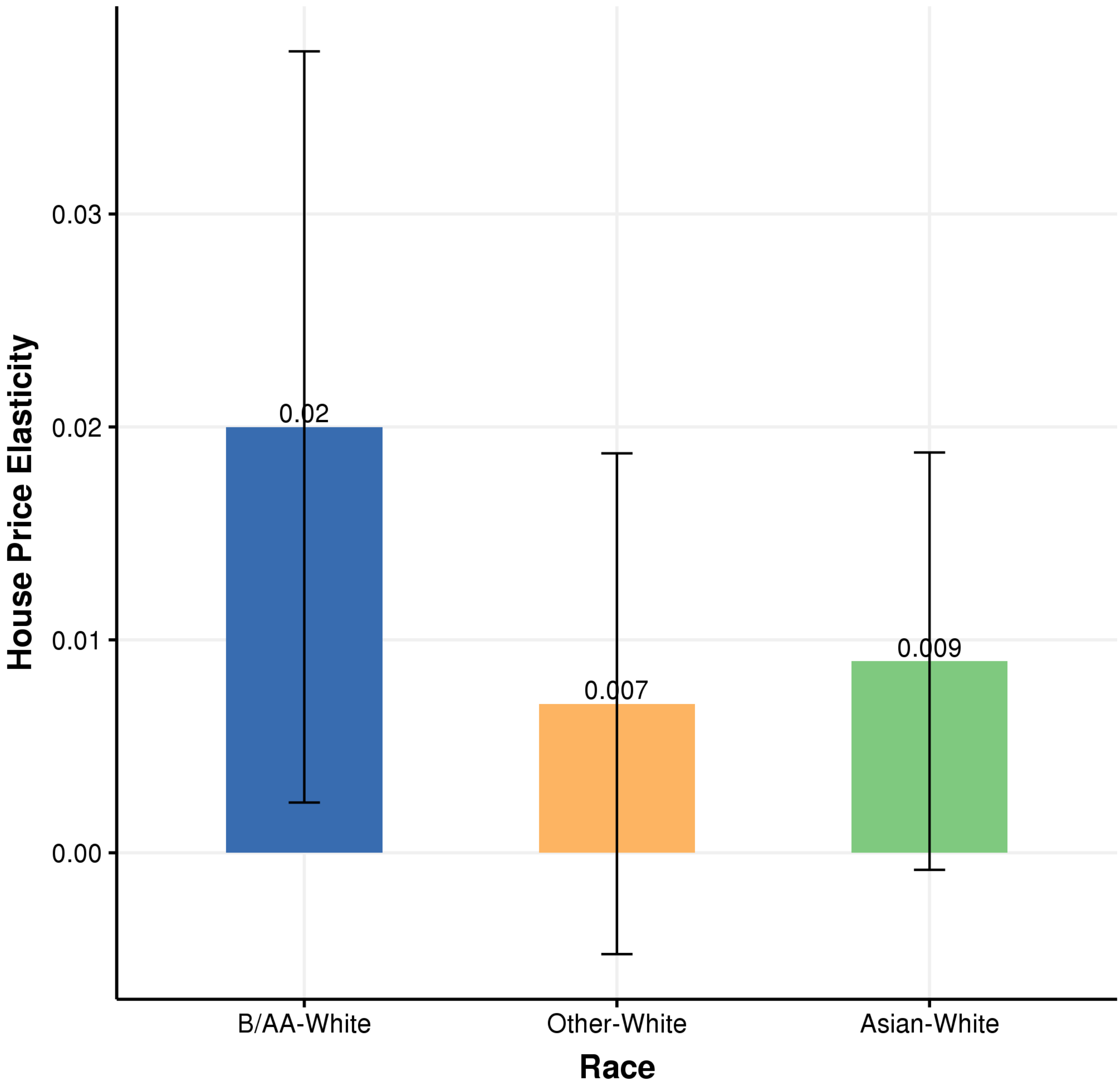


Figure 8. Map of Weighted Average Elasticities, County Level
This map shows the weighted average of the estimated household level elasticities for 2004 to 2016 at the county level. Each household's estimate is weighted by the survey weights and the elasticities come from the base GRF model.

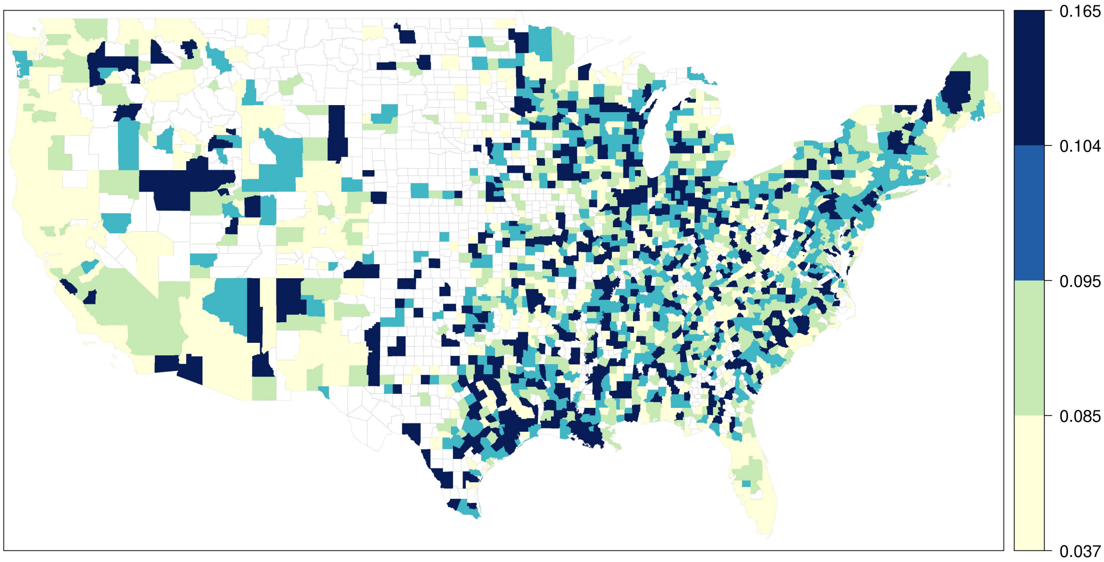


Figure 9. Map of Weighted Average Elasticities, State Level
This map shows the weighted average of the estimated household level elasticities for 2004 to 2016 at the state level. Each household's estimate is weighted by the survey weights and the elasticities come from the base GRF model.

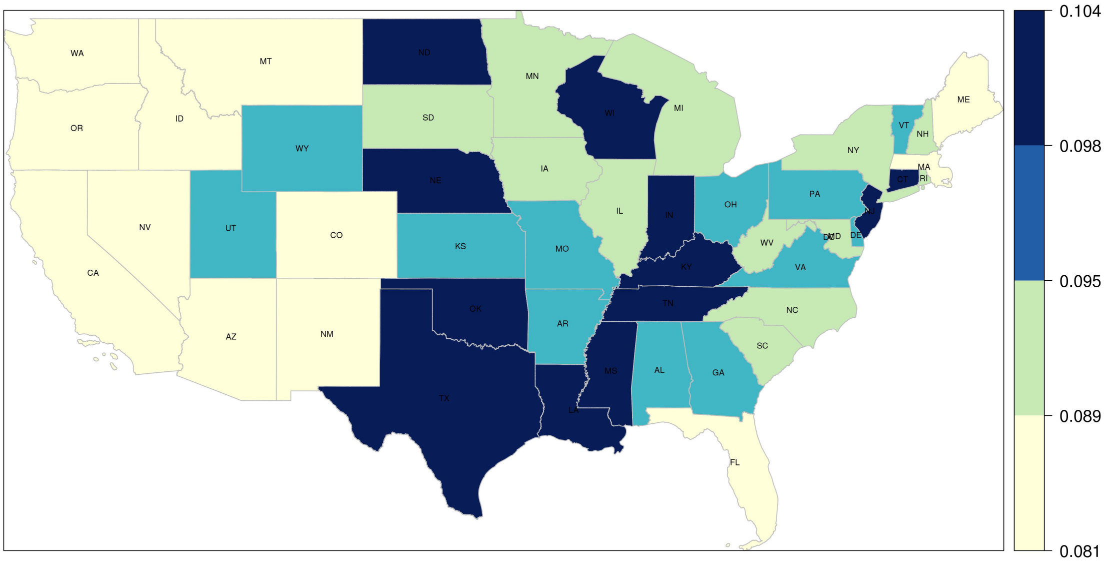


Figure 10. Covariance of Local House Price Growths and the House Price Elasticities
This plot shows the covariance between the house price growth and the household level elasticities estimated using the main GRF model at each year. House prices are normalized by the CPI of 2004 before growth rates are calculated.

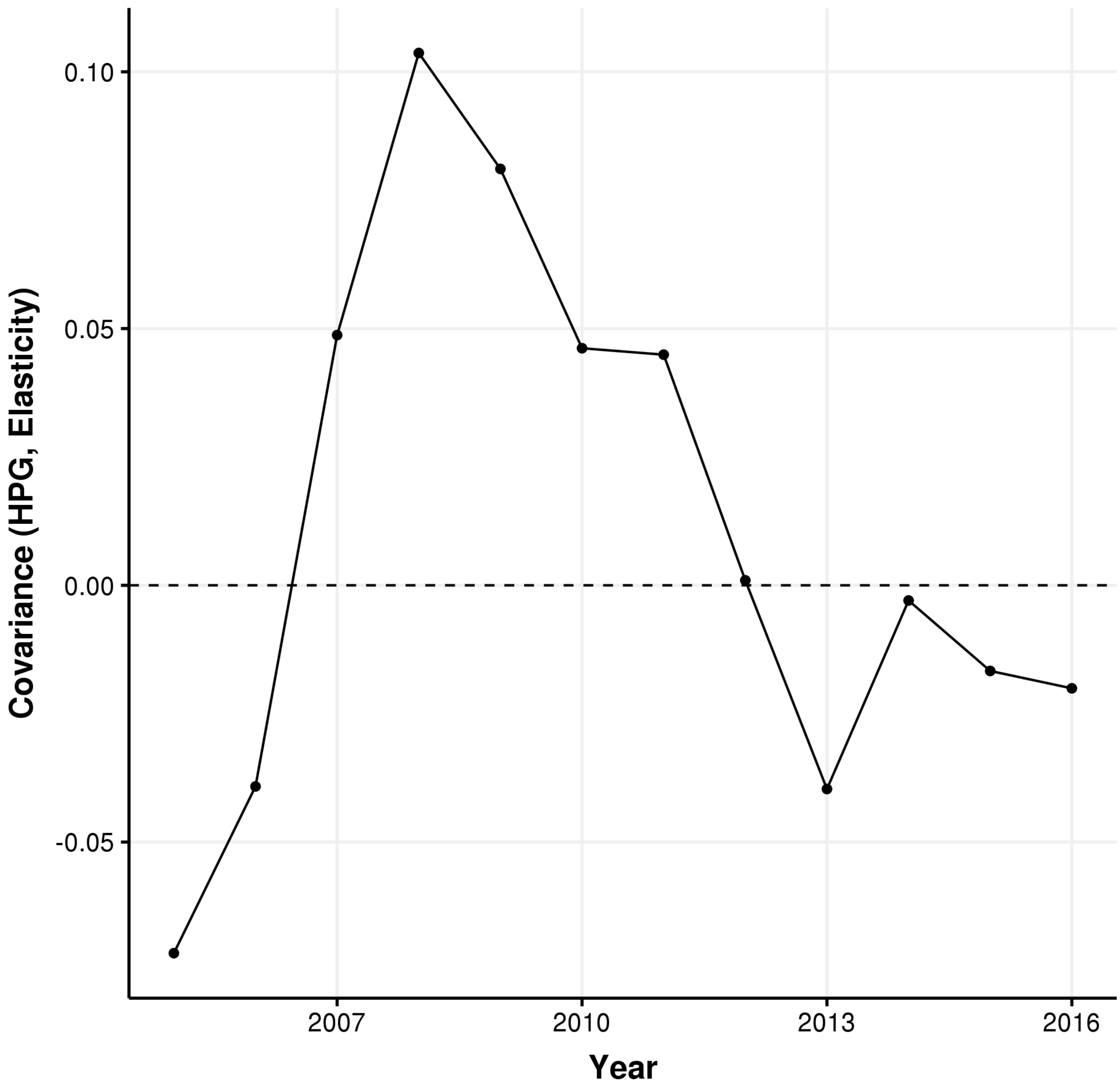


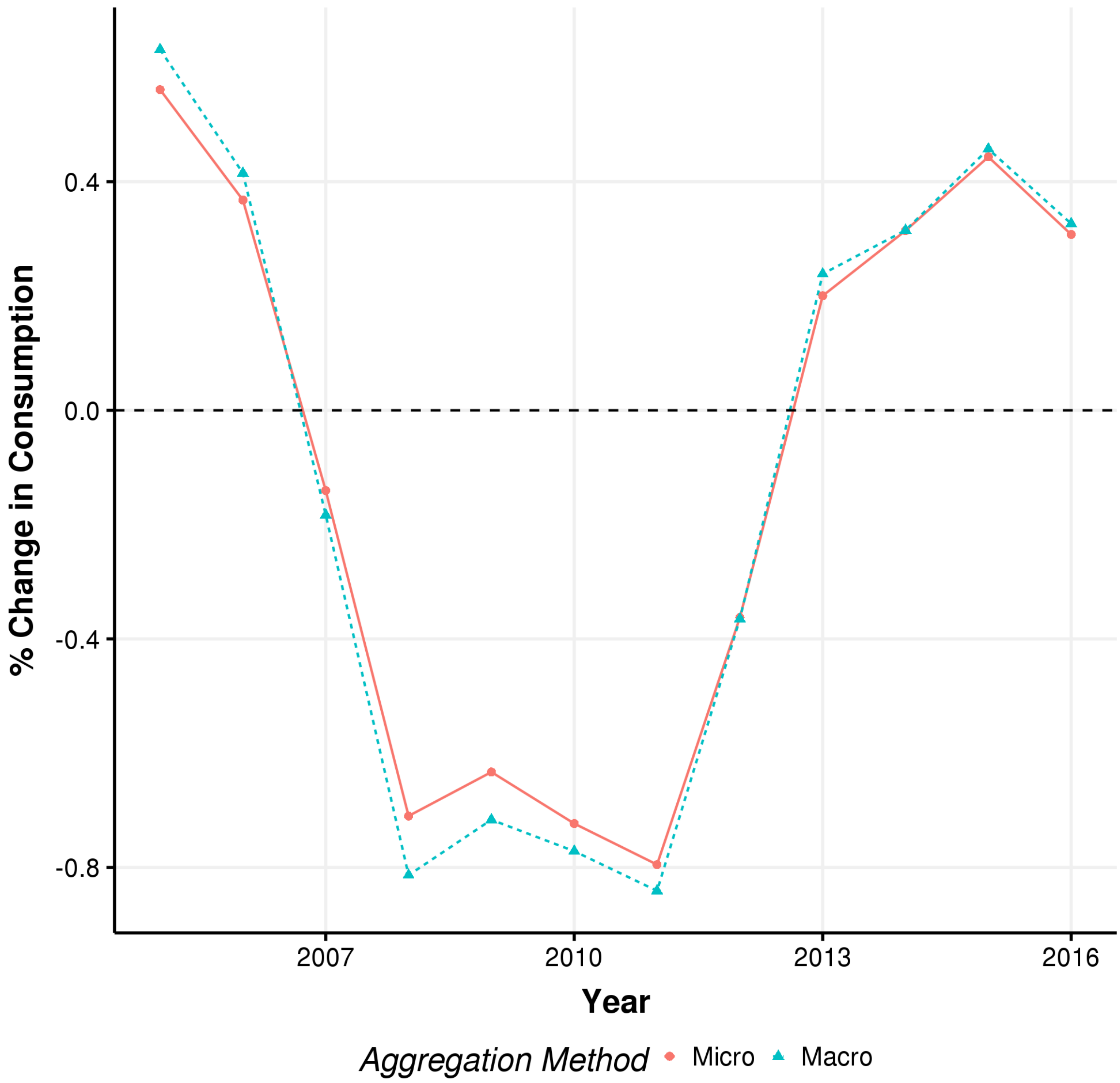
Figure 11.
In this plot, estimates at each year correspond to the following measures:

$$\Delta C_{Macro} = \bar{\theta} \sum_{c} \Delta hp_{ct}/N$$

and

$$\Delta C_{Micro} = \sum_{c} \hat{\theta}_{ct} \Delta hp_{ct}/N$$

Where ΔC is the log change in normalized consumption expenditure plotted in this figure, $\bar{\theta}$ is average elasticity for the entire country, N is total number of counties, $\hat{\theta}_{ct}$ is the elasticity at the county and year level. Δhp_{ct} shows the log change in normalized house price indexes at the county and year level. All aggregations are weighted by the survey weights. The table below this plot shows the corresponding values to each point depicted.



Year	2005	2006	2007	2008	2009	2010	2011	2012	2013	2014	2015	2016
Micro	0.6	0.4	0.1	0.7	0.6	0.7	0.8	0.4	0.2	0.3	0.4	0.3
Macro	0.7	0.4	0.2	0.9	0.7	0.6	0.7	0.3	0.2	0.4	0.4	0.4
Difference	0.2	0.005	0.1	0.2	0.1	0.1	0.1	0.1	0.03	0.1	0.04	0.1

Figure 12. Correlations between County Elasticity and House Price Growth, Years 2005 to 2008
Each sub-figure in this plot shows the correlation between the house price elasticity and house price growth at the county level. The county level elasticities are weighted averages of household level elasticities using the survey weights and the house price growth is the log change in normalized house price indexes. The linear regression line along with the standard error is shown. Observations (counties) are divided to above and below average variance based on the variance of house prices in that specific year.

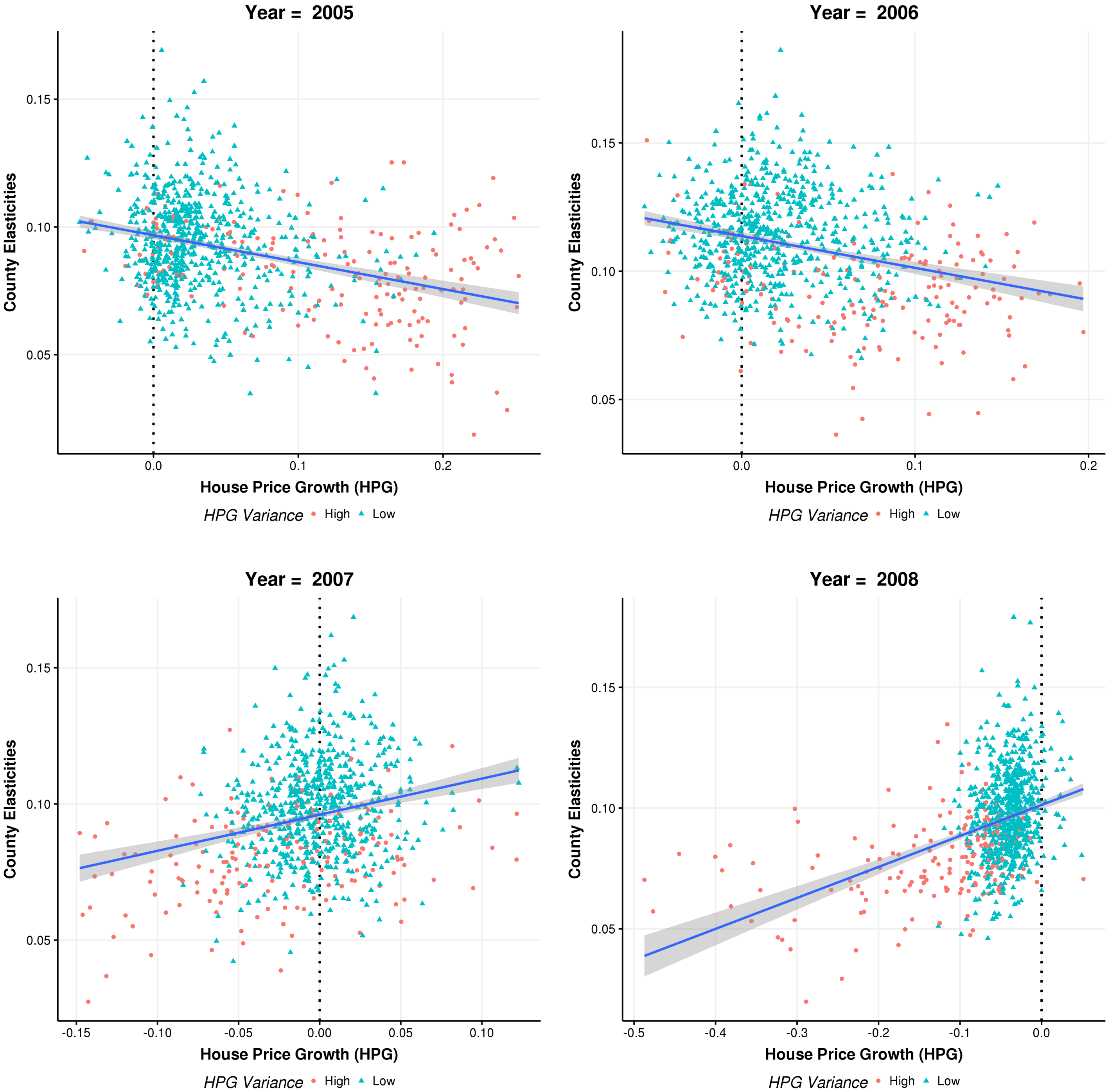


Figure 13. Correlations between County Elasticity and House Price Growth, Years 2009 to 2012

Each sub-figure in this plot shows the correlation between the house price elasticity and house price growth at the county level. The county level elasticities are weighted averages of household level elasticities using the survey weights and the house price growth is the log change in normalized house price indexes. The linear regression line along with the standard error is shown. Observations (counties) are divided to above and below average variance based on the variance of house prices in that specific year.

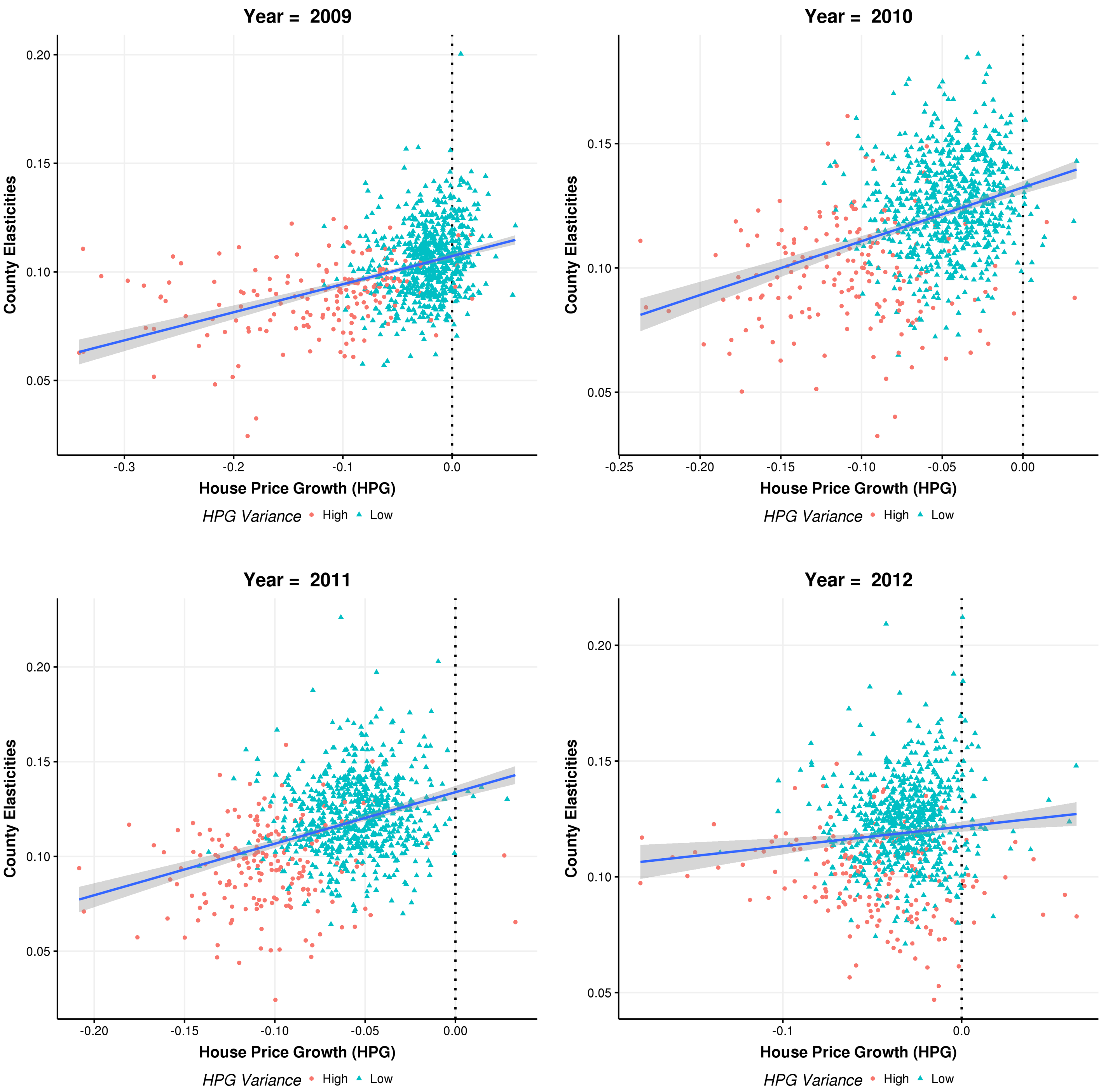
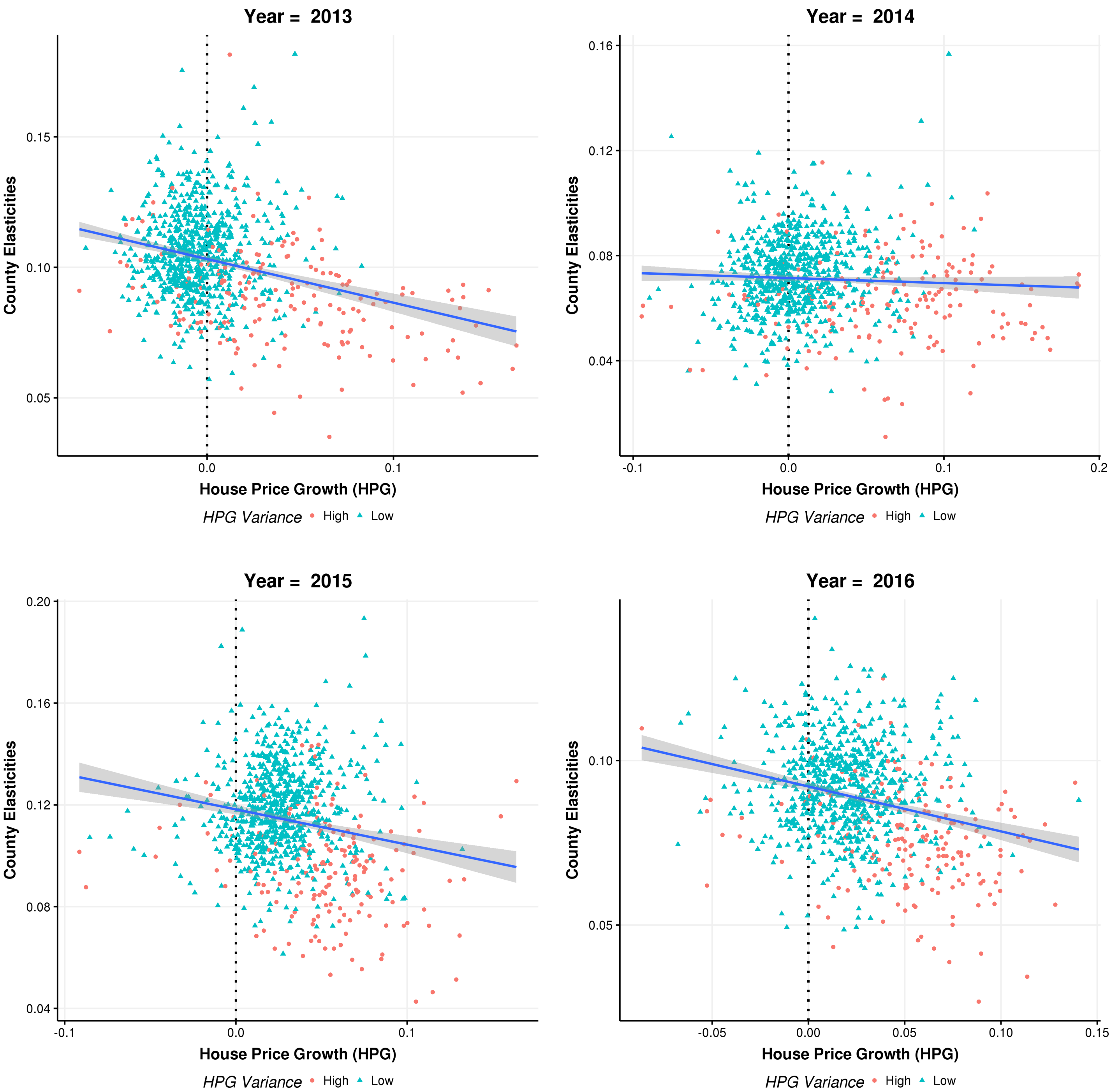


Figure 14. Correlations between County Elasticity and House Price Growth, Years 2013 to 2016

Each sub-figure in this plot shows the correlation between the house price elasticity and house price growth at the county level. The county level elasticities are weighted averages of household level elasticities using the survey weights and the house price growth is the log change in normalized house price indexes. The linear regression line along with the standard error is shown. Observations (counties) are divided to above and below average variance based on the variance of house prices in that specific year.



C Other Data

Other complimentary datasets are listed below.

- I deflate consumption and income and house prices by U.S. Bureau of Labor Statistics (2018)'s CPI for Urban areas, Not seasonally adjusted, base year of 2004 from U.S. Bureau of Labor Statistics (2018).
- Maps shape files have been extracted from Steven Manson, Jonathan Schroeder, David Van Riper, and Steven Ruggles (2019)
- *NBER.*— The county to CBSA cross walk is coming from NBER dataset available here.
- Debt to income the public version of county-level debt to income ratio using the Federal Reserve Bank of New York and IRS data for income taken from Mian et al. (2013) and available here.

D GRF Data Preparation

In the GRF model, I use the following control variables (the number in parenthesis shows the number of categories of each dummy variable): Age, income, education (3), employment categories (3), child under 18 (2), household size, marital status (2), race (4), residence type (5), year and county debt to income ratio from Mian and Sufi (2013). These are same control variables as in the OLS and IV regressions to increase comparability.

In the original data, household demographics are reported in more detailed categories which I further merge them into larger groups for two main reasons. One, to estimate more precise elasticities for each group. For instance, t is harder to find differences in behavior based on very fine categories of education but larger categories can show more variations.

The second reason is more technical. Forests only take numeric variables as inputs. Therefore, categorical variables need to be fed into the GRF model in the form of dummy variables. For instance, if variable A is available in 3 categories of a, b and c, one needs to create 3 different dummy variables and use them as inputs. Even if the categories are numeric[33], it is still necessary to create dummy variables.

In the following tables, I show the original and reduced categories for each categorical control variable that I use in the GRF model.

Education The data provides the education level of both heads of the household in 6 numeric categories. I round the average of the household heads education attainments and create new categories of the education level for each household. Then, I reduce them to 3 categories based on the following table. In case there is only one head of household, the education attainment of the head is used.

[33] Unless the differences between the numeric categories have meanings for the analysis, it is recommended to create dummy variables.

Original Categories	*Reduced Categories*
Grade School, Some High School, Graduated High School	High School
Some College, Graduated College	Undergraduate
Post College Grad	Graduate

Employment Households report their work hours in 3 categories of under 30 hours, 30 to 34 hours and above 35. I use this information to create the following 3 categories for the employment status of household heads: both heads employed (in case either there are two heads of household and both reported some hours worked or there is one head and is employed), one head employed (in case there are two heads and only one has reported hours worked) and both heads unemployed (in case there is only one head and he is unemployed or there are two heads and both are unemployed).

Child Under 18 One of the demographic variables reported by households is the age of their children if under 18. This variable is reported in 8 categories, which 7 of them are corresponding the age bins of the children and the last category is if they do not have any children below 18 years old. I compress this variable into a dummy variable of having a child under 18 or not.

Marital Status Marital status is reported in 8 categories which I then compress into a dummy variable of being married or not.

Race This variable is reported in 4 categories of White/Caucasian, Black/African American, Asian and Other. I include all 4 categories in the GRF model.

Residence Type Households report their type of dwelling in 7 categories which I then merge them into 5 detailed in the following table.

Education

Original Categories	*Reduced Categories*
One Family House, One Family House (Condo/Coop)	Single Family Dwelling
Two Family, Two Family House (Condo/Coop)	Undergraduate
Three+ Family House, Three+ Family House (Condo/Coop)	Two Family Dwelling
Mobile Home or Trailer	Mobile Home or Trailer

E Regression Trees, Bagging and Random Forest

In this sections, I provide an introductory description of Regression trees, Bagging and Random Forest models. However, readers are encouraged to look at chapters 9 and 10 of Hastie et al. (2013) available here.

F Generalized Random Forest Estimation

Gradient-Based Splitting Scheme

GRF replaces the kernel weighting with random forest adaptive weighting in order to estimate $\theta(x)$. In RF weighting, observations that are more frequently placed into the same leaf with observation i receive higher weights. The observations that receive higher weights are closer neighbors to the observation at hand.

The weighting idea is similar to the weights in local linear regressions, the goal is to weight more heavily the households (observations) that are near the target household (observation i) and less heavily the ones that are further away.

Tuning and Out of Bag Goodness of Fit Estimates

I can estimate the out of bad goodness of fit (GOF) which gives a measure of model performance with a low computational cost[34]. For each observation, the algorithm selects the trees that have *not* involved the observation at hand in the training process. These trees are then used to estimate the error of prediction for that particular observation as a measure of GOF.

The general convention is not to very carefully tune the RF parameters or not tune them at all. The argument is that forests do not overfit since they are self-regularizing and the final outcome is the average of many models. However, as Athey et al. (2019b) show the performance of the model improves with tuning[35]. The forest parameters that I tune include number of trees, subsample, min node size (deepness) and the subset of the covariates.Besides, they use honest estimation[36] which further decreases the possibility of overfitting.

G GRF Parameters and Tuning

This section describes the parameters and procedures used to tune the GRF model. The following is the list of the parameters used and their interpretations.

- Number of variables considered at each split = 2,

  This number controls the complexity of the tree and overfitting. At each node two variables are randomly selected and the one which can reduce the error the most and satisfy the minimum node size requirement is selected.

- Minimum node size = 77

  This hyperparameter requires each final node to have at least 77 observations. Increasing the minimum node size results in smaller trees and prevents over fitting. However, a very high minimum node size can lead to a stump and underfitting.

[34] One does not need to re-run the forest (or do bootstrap) in order to obtain a measure of model performance. The idea of using out of bag observations allows me to use the originally trained model to estimate GOF and hence run the forest only once.

[35] They have included the tuning options in their packages as well.

[36] In honest estimation, one uses a part of the training data for training and the other part for tuning. In other words, the tuning is done only on a subset of the training data and not all of it. In short, tuning on a different sample that the model was trained on, will help prevent overfitting.

- $\alpha = 0.05$

  Controls the maximum imbalance of a split in terms of control and treatment. The higher this hyperparameter is set the more strict the tree is to have a more balanced set of treatment and control observations at each node.

- Number of trees = 1000

  This is the number of trees used in the forest. One could potentially increase the number of trees and reduce the standard errors of predictions, however there is a computational cost for increasing the iterations.

- Sample fraction = 0.5

  Fraction of data used to grow the trees. The rest of the data is used to draw predictions.

- CI group size = 30

  Controls the number of trees in each group for variance estimation

- Clusters = CBSA

  Takes the cluster into account for sampling by giving more weights to observations within the cluster. This features works very similar to the fixed effect notion in linear regression setting.

As Wager[37] mentions "optimal methods for tuning heterogeneous treatment effect estimators with instrumental variables (IVs) are still an open research topic". I selected the parameters to reduce the variance of estimations and also to be able to replicate the linear IV regression results on average. I have also tested my tuning procedure by tuning the first and second stage separately using their proposed method and found similar results.

H GRF Robustness Checks

To test the robustness of the main GRF model, I train the model using the same parameters and variables on a random sample of 80% of the data, call this model1 and the original model trained on the entire sample model0. In the next step, I use the 20% of the data that has not been used in training model1 to predict the household level elasticities using both models. Finally, I estimate the mean square error of prediction, assuming that the outcome of the original model (model0) is the truth. I then repeat this procedure 50 times and calculate the average mean squared error. The steps are summarized bellow:

- Train the model on a random sample of 80% of the data, call it model1
- Predict the elasticities using model1 and the remaining 20% of the data, call it $\hat{\theta}_1$
- Predict the elasticities using model0 (the original model) and the remaining 20% of the data, call it $\hat{\theta}_0$

[37]He is an assistant professor of Operations, Information, and Technology at the Stanford Graduate School of Business and one of the authors of Athey et al. (2019b) paper. The quote is taken from GRF's git hub discussions.

- Calculate the root mean squared error of prediction assuming that $\hat{\theta}_0$ is the true parameter using the following formula

$$mse_s = \sqrt{\hat{\theta}_1 - \hat{\theta}_0}$$

- Repeat 50 times and calculate the average of the root mean squared error

$$mse = \sum_{s=1}^{S} mse_s / S$$