NBER WORKING PAPER SERIES

HOUSING DEMAND, COST-OF-LIVING INEQUALITY, AND THE AFFORDABILITY CRISIS

David Albouy
Gabriel Ehrlich
Yingyi Liu

Working Paper 22816
http://www.nber.org/papers/w22816

NATIONAL BUREAU OF ECONOMIC RESEARCH
1050 Massachusetts Avenue
Cambridge, MA 02138
November 2016

During work on this project, Albouy was a David C. Lincoln Fellow for the Lincoln Institute of Land Policy. Any mistakes are our own. We thank seminar participants at the Budapest Institute of Economics, Dalhousie, Homer Hoyt Institute, LSE (CEPR), Monash, National University of Singapore, New Economic School (Moscow), Tinbergen Institute, University of Bern, UIC, UNSW, USC (Price). We also thank conference participants at the AREUEA, ASSA, and SEA annual meetings; CHUM-Chicago Fed; IEB IV Workshop on Urban Economics; NBER Conference on International Comparisons of Income, Prices, and Production; the Ziman-UCLA Conference on Housing Affordability; David Card, Tom Davidoff, Angus Deaton, Jessie Handbury, Ed Olsen, Chien-Wen Peng, Changcheng Song , Lewis Warren, and Jing Wu for their comments and input. The views expressed herein are those of the authors and do not necessarily reflect the views of the National Bureau of Economic Research.

Housing Demand, Cost-of-Living Inequality, and the Affordability Crisis
David Albouy, Gabriel Ehrlich, and Yingyi Liu
NBER Working Paper No. 22816
November 2016
JEL No. D12,E31,R21

ABSTRACT

Since 1970, housing's relative price, share of expenditure, and ``unaffordability" have all grown. We estimate housing demand using a novel compensated framework over space and an uncompensated framework over time. Our specifications pass tests imposed by rationality and household mobility. Housing demand is income and price inelastic, and appears to fall with household size. We provide a numerical non-homothetic constant elasticity of substitution utility function for improved quantitative modeling. An ideal cost-of-living index demonstrates that the poor have been disproportionately impacted by rising relative rents, which have greatly amplified increases in real income inequality.

David Albouy
Department of Economics
University of Illinois at Urbana-Champaign
214 David Kinley Hall
Urbana, IL 61801-3606
and NBER
albouy@illinois.edu

Gabriel Ehrlich
Department of Economics
University of Michigan
611 Tappan St
Ann Arbor, MI 48109-1220
gehrlich@umich.edu

Yingyi Liu
Dept of Economics
University of Illinois
1407 W. Gregory Dr
Urbana, IL 61801
liu195@illinois.edu

1 Introduction

Food, clothing, and shelter are all considered to be basic needs. Yet, as Figure 1A shows, from 1959 to 2014 the proportion of personal consumption expenditures in the United States devoted to food and clothing fell from 27.4 percent to 10.6 percent, while the fraction devoted to housing and utilities rose from 16.1 to 18.1 percent. Data from the American Housing Survey and Consumer Expenditure Survey indicate more dramatic increases in housing's share: about 7 percentage points since 1970, as illustrated in Figure 1B. This growth has been even sharper among renting households. Figure 1C shows that the percentage of renting households facing "moderate" or "extreme" affordability burdens, defined as spending more than 30 or 50 percent of their income on housing, has risen by 20 and 15 percentage points, respectively. Meanwhile, the home-ownership rate has not seen a persistent rise. These trends support the Secretary of Housing and Urban Development's recent claim: "We are in the midst of the worst rental affordability crisis that this country has known" (Olick 2013).[1] As mapped in figure 2, this crisis is particularly severe in in large cities and along the coasts.

The increasing share of expenditures on housing appears to contradict the traditional view that housing is a necessity. On that view, the expenditure share on housing should have fallen as average incomes have risen over time. Another common assumption, that preferences over housing relative to other goods are unit elastic in price and income (i.e., Cobb-Douglas), is also incapable of explaining increases in un-affordability measures or their variation over space. Below, we consider preferences and changes in the economy that explain both the rise in the housing expenditure share over time, as well as the spatial variation in that share.

First and most important is that the price of housing relative to other goods has risen substantially. Figure 3A shows that the price of housing (or shelter) services has risen almost 40 percent relative to other goods since 1970, as measured by the Bureau of Labor Statistics' Consumer Price

[1] The Joint Center for Housing Studies of Harvard University (JCHS, 2013) documents that from 2000 to 2012, the median share of renters' incomes devoted to contract rent rose nearly five percentage points to 27.4 percent, and that 28 percent of renting households now spend more than half of their incomes on rent. Glaeser and Gyourko (2008) present a thoughtful critique of affordability indices as well as a discussion of how rental expenditures have changed over time.

Index (CPI). Critiques of the CPI suggest that the official series understates increases in the relative price of shelter severely, while it overstates increases in the absolute prices of most other goods and services. When we adopt revisions proposed by Crone, Nakamura and Voith (2010) and Boskin (1996), the relative price of shelter has increased by 120 percent since 1970. If housing demand is price inelastic, the large relative price increase of housing can help to explain its rising share.

Second, rising income inequality has lowered median incomes relative to mean incomes, as seen in Figure 3B. This phenomenon is particularly acute among renters, whose incomes have seen weaker gains relative to homeowners. Thus, rising incomes have had smaller effects for the majority of renters. Third, falling household sizes may have increased demand. Figure 3C shows that households have shrunk in size by almost 30 percent. This has reduced economies of scale in housing consumption, thereby increasing per-capita demand for housing.[2]

To investigate these issues, we estimate housing demand using an intuitive framework motivated by spatial equilibrium conditions. Section 3 demonstrates that cross-sectional data lends itself to estimating compensated (Hicksian) housing demand functions, as mobility equalizes the utility households receive from living in different locations. On the other hand, time-series data lends itself to uncompensated (Marshallian) demand functions, as utility levels may change over time. An important innovation of our approach is that we use data on non-housing prices to test restrictions imposed by demand theory, thus checking the validity of our specifications. This provides an unconventional examination of demand theory through spatial variation, rather than more conventional temporal variation (e.g. Deaton 1986, Blundell et al. 1993). We also estimate household economies of scale in the spirit of Barten (1964), and develop methods for estimating the direct effect of amenities on housing demand.

Under such restrictions, we integrate a demand equation into a utility function in the non-homothetic constant elasticity-of-substitution (NH-CES) framework. This function should be use-

[2]This increased demand requires that housing is indeed price inelastic. Of course, household sizes may be responsive to incomes and prices. We leave the topic of endogenous household sizes to future work. Note the proportion of children has fallen as the population has aged. Children under 18 accounted for nearly 24 percent of all household members in 1970, but accounted for only 15 percent as of 2013. To the extent that adults desire more housing than children do, this trend should have also raised housing expenditures.

ful to researchers modeling housing consumption and provides cost-of-living indices across space and time for different income, household sizes, and amenity levels, which account for income and substitution effects.

Our compensated estimates suggest that the uncompensated own-price, income, and substitution elasticities are all near two-thirds in absolute value. Our tests suggest that estimates based on cross-metropolitan (as opposed to within-metropolitan) variation are likely unaffected by household sorting, which in principle would bias the estimates away from zero. The quality of our control for utility, based on the wage predicted by worker skills, cannot be fully tested, but likely suffers from fewer biases than other measures.

We find some evidence of economies of scale in housing, in proportion to the square-root of housing size. A 30-percent reduction in residents per household reduces household housing consumption by only 15 percent, increasing per-capita consumption by the same percentage. Furthermore, estimates suggest that hilliness and hot weather may increase housing demand independently of prices.

Time-series patterns are largely consistent with our cross-sectional results. With the adjusted CPI, an elasticity of substitution slightly lower, or an income elasticity slightly higher than our estimates — in the direction of our expected biases — would completely explain housing's rising share in national expenditure. Demographic effects reinforce these trends. Furthermore, rising rents and growing income equality explain most of the "affordability crisis" affecting renters. There may be some residual increases in the housing share that warrants additional exploration.

We demonstrate numerically how an ideal cost-of-living index varies non-linearly with income and prices across space and time. While the prices of many goods have become much cheaper over time, housing rents do not appear to have made nearly as much progress. By differentially impacting the poor, increases in the relative price of housing have increased real income inequality by 25 percent since 1970.

2 Motivation and Related Literature

Our explanation of competing income and relative-price effects is illustrated in Figure 4, with production possibility frontiers (PPF) and indifference curves for housing and non-housing goods. Over time, the PPF has expanded further in the direction of non-housing goods: these goods are traded internationally and are subject to greater technological improvements. With this expansion, both income effects (illustrated from point A to point B) and relative-price effects (from B to C) lead households to increase their consumption of non-housing goods more than of housing. The income effect causes housing's share to fall (compare B and D), but the rise in the relative price causes housing's share to rise if the substitution response is limited (compare C and E).

Previous researchers have estimated a wide range of price and income elasticities. Articles reviewed in Mayo (1981) find uncompensated price elasticities from slightly positive to less than minus one. Estimates closest to ours include Pollinsky and Ellwood's (1979) estimate of -0.7 and Hanushek and Quigley's (1980) experimental estimates of -0.64 in Pittsburgh and -0.45 in Phoenix.[3] Classical studies such as Engel (1857) and Schwabe (1868) estimated the income (or more precisely, expenditure) elasticity of housing demand to be less than one, which became known as "Schwabe's Law of Rent".[4] As discussed in Mayo (1981), a source of contention is how to measure income: most use a proxy measure of "permanent income" to correct for attenuation bias caused by transitory income. Davis and Ortalo-Magne (2011) argue that the median expenditure share on rent across metros — the evidence we focus on — is roughly constant across metro areas, consistent with price and income elasticities of one.

Modelers have taken great latitude in interpreting such disparate findings. Indeed, housing

[3] Other articles include Muth (1960), Reid (1962), Rosen (1985), Goodman and Kawai (1986), Goodman (1988) Ermisch et al. (1996), Goodman (2002), and Ioannides and Zabel (2003). Most estimate uncompensated price elasticities ranging from -1 to -0.3 and income elasticities from 0.4 to 1. While some studies use non-housing price data to deflate their numbers, none use it to test the validity of the housing demand specification, as we do here. Few articles estimate elastic price demand, with elasticities greater than one. Kau and Sirmans (1979) estimated price elasticity shifting from -2.25 to -1 from year 1876 to 1970 using historical data from Chicago. However, these are based off land-price gradients and are not robust to expected sorting behaviors described below.

[4] See Stigler (1954) for a discussion. As summarized by De Leeuw (1971), Mayo (1981) and later Harmon (1988), some studies are inconsistent with Schwabe's Law. See Hansen et al. (1998) and references therein for estimates less than one, and Muth (1960) and Cheshire and Sheppard (1998) for an estimate greater than one, noting that the latter study estimates elasticities for housing attributes rather than for a unified bundle.

demand is key to understanding house prices, tax incidence, population density, and location decisions. Many assume a fixed demand for housing, perfectly inelastic to price and income. This provides a simple derivation of the mono-centric city model, seen from Mills (1967) to Desmet and Rossi-Hansberg (2013). Other models, such as the search and matching model of Piazzesi and Schneider (2010), assume housing demand is inelastic to income but not to prices. Unit-elastic demand — derived from Cobb-Douglas preferences — is especially common: examples include Eeckhout (2004), Michaels, Rauch and Redding (2012), and Guerreiri, Hartley and Hurst (2013). While abstraction is often necessary, these disparate assumptions make it hard to reconcile different findings, and may lead to incorrect conclusions in some contexts.

Indeed, the issue of housing "affordability," especially as measured by high expenditure shares, makes the most sense when demand is income- and price-inelastic. This is especially true for low-income households, whose incomes have lagged particularly in America's largest, most expensive cities (Baum-Snow and Pavan, 2013). Low-skilled workers' greater housing expenditure share explains their choosing to live in cheaper cities (Moretti 2013), while those remaining in expensive cities must earn higher wage premia to do so relative to the premia required by more-skilled workers (Black, Kolesnikova, and Taylor 2009).[5]

The secular rise in housing expenditures appears to be understudied. It is in line with Piketty's (2014) finding that the value of residential capital relative to output rose substantially over the last century.[6] With inelastic demand for land in both consumption and production, land's value can take up an increasing share of the economy, reviving fears of Ricardo (1817) and George (1879).[7]

[5]Handbury (2013) estimates a non-homothetic log-logit utility function with a CES superstructure to argue that high-income households find large cities to be more "affordable" by containing a greater range of groceries suited to their tastes. We find that large cities are more affordable for high-income households as they spend less on housing.

[6]Gyourko, Sinai, and Mayer (2013) find housing values' differences between typical and highest-price locations widened considerably since 1960. Rognlie (2015) shows that the postwar increase in the share of income flowing to capital is largely concentrated in the housing sector, which La Cava (2016) shows is mainly due to higher imputed rental income to owner-occupiers. Davis and Heathcote (2007) present evidence of persistent real growth in land values, accounting for an increasing share of housing values. This evidence is consistent with limited substitution between land and non-land inputs in housing production, as found in Albouy and Ehrlich (2012). We note that housing is a capital asset that provides flow consumption services to its owner. This asset is a composite of land and structure, the latter of which typically depreciates over time. We follow the bulk of the literature in estimating demand for a composite housing good, but the shape of the housing demand function can have important implications for land values separately from housing values.

[7]This may happen if land-saving technological improvements are weak or stifled by regulation. Thus, rising de-

3 Housing Demand as Prices, Incomes, and Amenities Vary

Here we present a standard static model of housing demand embedded in a spatial equilibrium framework with local household amenities, similar to the settings of Rosen (1979), Roback (1982), and Albouy (2016). We derive expressions for the share of expenditure devoted to housing, which motivate regressions to identify the parameters of the compensated and uncompensated housing demand functions, amended to allow for variation in household size. We then construct utility and cost-of-living functions that allow for imperfect substitution, non-homotheticity, varying household sizes, and variation in amenities.

3.1 Household Budgets and Preferences

The national economy contains many cities, indexed by j, which share a population of mobile households, who supply one unit of labor where they live. They consume a housing good y with price p^j, and a non-housing good x with price c^j.[8] Households earn total income $m^j = I+(1-\tau)w^j$, determined by unearned income, I, which does not vary by city, and local wage levels w^j, after taxes, τ.[9] Household preferences over the consumption good, housing, and location are modeled by a utility function $U(x, y; Q^j)$, where Q^j represents a city-specific amenity bundle conceptualized as "quality-of-life". The indirect utility function for a household in city j is then given by $V(p^j, c^j, m^j; Q^j) = \max_{x,y}(U(x, y; Q^j)|c^j x + p^j y = (1-\tau)w^j + I)$. The expenditure function for a household in city j is likewise given by $e(p^j, c^j, u; Q^j) = \min_{x,y}(c^j x + p^j y|U(x, y; Q^j) \geq u)$.

3.2 The Housing Expenditure Share and Uncompensated Demand

To take the model to the data, we approximate the relationships described above around their national average values. Denote the fraction of household expenditures on housing in city j as

mand may reverse earlier declines in land values engendered by transportation improvements.

[8] For simplicity, the exposition of the theoretical model will refer to a system of cities and call individual geographical units as such. However, the empirical work using the Consumer Expenditure Survey (CEX) data will be at a partly-disaggregated state level. Therefore, the geographies considered in this model are more properly considered 'areas', with the term 'city' used for concreteness.

[9] In this static setting, household expenditure equals household income.

$s_y^j \equiv (p^j y^j)/m^j$. Log-linearizing this equation produces the identity $\hat{s}_y^j = \hat{p}^j + \hat{y}^j - \hat{m}^j$.[10] We assume that households take local price and income levels as given, so that the only behavioral variable in s_y is housing consumption, y, which is determined by the uncompensated (Marshallian) demand function $y^j = y(p^j, c^j, m^j; Q^j)$. Log-linearizing demand, $\hat{y}^j \equiv \epsilon_{y,p}\hat{p}^j + \epsilon_{y,c}\hat{c}^j + \epsilon_{y,m}\hat{m}^j + \epsilon_{y,Q}\hat{Q}^j$. The parameter $\epsilon_{y,p}$ is the uncompensated own-price elasticity of housing demand, $\epsilon_{y,c}$ is the uncompensated cross-price elasticity, $\epsilon_{y,m}$ is the income elasticity, and $\epsilon_{y,Q}$ is the amenity elaticity amenities. If housing is a normal good, then $\epsilon_{y,m} > 0$, and housing obeys the law of demand that $\epsilon_{y,p} < 0$. It is *a priori* unclear whether housing is a gross substitute for non-housing goods, i.e., whether $\epsilon_{y,c} > 0$, because the cross-price elasticity exhibits a positive substitution effect and a negative income effect, each of unknown magnitudes. Housing may be a gross complement or substitute for amenities, i.e. $\epsilon_{y,Q} \gtreqless 0$, if amenities alter the marginal rate of substitution between housing and non-housing goods.

Combining the identities above demonstrates how the housing share depends on local variables:

$$\hat{s}_y^j = (1 + \epsilon_{y,p})\hat{p}^j + \epsilon_{y,c}\hat{c}^j + (\epsilon_{y,m} - 1)\hat{m}^j + \epsilon_{y,Q}\hat{Q}^j \tag{1}$$

Unrestricted, equation (1) is merely definitional. Rationality of preferences restricts the demand function to be homogenous of degree zero in prices and income (p, c, m), so that $\epsilon_{y,p} + \epsilon_{y,c} + \epsilon_{y,m} = 0$. This restriction of "no money illusion" requires that proportional increases in all prices and income do not lead to changes in behavior.[11]

Adding a constant and error term to equation (1) motivates these regression equations:

$$\ln s_y^j = \alpha_0 + \alpha_1 \ln p^j + \alpha_2 \ln c^j + \alpha_3 \ln m^j + \alpha_4 q^j + e^j \tag{2a}$$

$$= \alpha_0 + \alpha_1(\ln p^j - \ln c^j) + \alpha_3(\ln m^j - \ln c^j) + \alpha_4 q^j + e^j \tag{2b}$$

[10] A hat over a variable represents its log deviation from the (geometric) national average, i.e., $\hat{z}^j = d \ln z^j = dz^j/\bar{z}$.

[11] We do not model how households with low tastes for housing may be inclined to seek out more amenable areas (see Black et al. 2002). Albouy and Lue (2015) present evidence that household sizes, age, and marital status vary little across metropolitan areas (they vary more within), suggesting such selection issues are not of first-order importance. While discussed more below, our rationality test appears to rule out large amounts of sorting.

Equation (2b) follows from (2a) as homogeneity requires $\alpha_1+\alpha_2+\alpha_3 = 0$. If we subtract the means of the right-hand side variables, the regression coefficients are related to the demand parameters as: $\alpha_0 = \ln \bar{s_y}$, $\alpha_1 = 1 + \epsilon_{y,p}$, $\alpha_2 = \epsilon_{y,c}$, and $\alpha_3 = \epsilon_{y,m} - 1$. $\bar{s_y} = e^{\alpha_0}$ is the geometric mean of expenditure shares. The own-price uncompensated elasticity is the coefficient on housing prices minus one, $\epsilon_{y,p} = \alpha_1 - 1$; income elasticity is the coefficient on income plus one, $\epsilon_{y,m} = \alpha_3 + 1$.

Quality of life is not observed directly but is proxied by observable amenities, q_j. Moreover, we model $\epsilon_{y,Q}$ as a vector because differing amenities may shift housing demand differently.[12]

Consistent estimation of equation (2a) requires that non-housing goods are properly accounted for by the index c^j, that preferences across cities are the same, that preferences can be aggregated, and that we have an appropriate (arguably permanent) measure of income m^j.

3.3 Compensated Demand with Household Mobility and Heterogeneity

The uncompensated demand function is converted into a compensated (Hicksian) demand function by substituting in the expenditure function, i.e. $y^H(p, c, m; Q) = y(p, c, e(p, c, u; Q); Q)$. Log-linearizing the expenditure function directly yields the mobility condition that local incomes will compensate for local prices and quality of life, conditional on utility:

$$\hat{m}^j = \bar{s_y}\hat{p}^j + (1 - \bar{s_y})\hat{c}^j + \epsilon_{m,Q}\hat{Q}^j + \epsilon_{m,u}\hat{u}^j \tag{3}$$

where $\epsilon_{m,u}$ is the elasticity of expenditures with respect to utility, and $\epsilon_{m,Q}$ is the elasticity of expenditures with respect to quality of life.

Substituting equation (3) into equation (1) and simplifying by the Slutsky equations gives the following relationships among the uncompensated (Marshallian) and compensated (Hicksian) price elasticities: $\epsilon_{y,p} = \epsilon^H_{y,p} - \bar{s_y}\epsilon_{y,m}$ and $\epsilon_{y,c} = \epsilon^H_{y,c} - \bar{s_x}\epsilon_{y,m}$. Here $\epsilon^H_{y,p}$ and $\epsilon^H_{y,c}$ are the compensated elasticities of housing demand with respect to housing and non-housing prices, respectively.[13] Ra-

[12] A priori we are unsure, nice climates could induce households to spend more of their time on the properties or away from them. Zivin and Neidell (2014) estimate the extent to which extreme heat and cold pushes people indoors.

[13] The first substitution yields $\hat{s}^j_y = (1 + \epsilon_{y,p} - \bar{s_y} + \bar{s_y}\epsilon_{y,m})\hat{p}^j + [\epsilon_{y,c} - (1 - \epsilon_{y,m})(1 - \bar{s_y})]\hat{c}^j + (\epsilon_{y,Q} - (1 - \epsilon_{y,m})\epsilon_{m,Q})\hat{Q}^j - (1 - \epsilon_{y,m})\epsilon_{m,u}\hat{u}^j$. Besides the Slutsky equations we also substitute in the identities $\epsilon^H_{y,Q} = \epsilon_{y,Q} +$

tionality requires that compensated demand functions are homogeneous of degree zero in prices, implying the own and cross-price elasticities sum to zero, $\epsilon_{y,p}^{H} + \epsilon_{y,c}^{H} = 0$.

Combining these insights yields the following equation for differences in the expenditure share in terms of relative prices, quality of life, and utility:

$$\hat{s}_y^j = (\epsilon_{y,p}^H + 1 - \bar{s_y})(\hat{p}^j - \hat{c}^j) + (\epsilon_{y,u}^H - \epsilon_{m,u})\hat{u}^j + (\epsilon_{y,Q}^H - \epsilon_{m,Q})\hat{Q}^j \tag{4}$$

Here $\epsilon_{y,Q}^H$ is the compensated elasticity of housing demand with respect to quality of life and $\epsilon_{y,u}^H$ is a similar elasticity for income.

We assume that similarly-skilled households are equally well-off across cities. When households are mobile, households should be indifferent across locations they inhabit, and utility by type of household will not vary across cities. Rather, utility differences will represent inherent differences across households, such as different earnings potentials. We parameterize income in city j as $m^j = \zeta^j w^j$, where ζ^j is an index of wage-earning skills, and w^j is the city-wide wage level that compensates households for living in that city.[14]

To interpret the coefficient, we posit that our utility function is money metric around national averages: $u(x, y; Q) = e(\bar{p}, \bar{c}, \tilde{u}(x, y; Q), Q)$. This normalization allows us to write utility differences in terms of differences in the skill index $\hat{u}^j = \hat{\zeta}^j$, and impose $\epsilon_{m,u} = 1$ and $\epsilon_{y,u}^H = \epsilon_{y,m}$.[15] This motivates a compensated empirical model that replaces income in (2a) with a skill index:

$$\ln s_y^j = \beta_0 + \beta_1 \hat{p}^j + \beta_2 \hat{c}^j + \beta_3 \hat{\zeta}^j + \beta_4 q^j + e^j \tag{5a}$$

$$= \beta_0 + \beta_1 (\hat{p}^j - \hat{c}^j) + \beta_3 \hat{\zeta}^j + \beta_4 q^j + e^j \tag{5b}$$

where $\beta_0 = \ln \bar{s_y}$, $\beta_1 = \epsilon_{y,p}^H + 1 - s_y = -\beta_2$ and $\beta_3 = \epsilon_{y,m} - 1$. In practice, $\hat{\zeta}^j$ is an index estimated

$\epsilon_{y,m}\epsilon_{m,Q}$ and $\epsilon_{y,u}^H = \epsilon_{y,m}\epsilon_{m,u}$ to get the resulting equation.

[14] When household types vary within city, the compensating wage differences will vary according to their tastes for housing, quality of life, and taxes.

[15] Note that we implicitly impose the restriction that the skill index affects housing consumption through income, and not through differences in tastes. If households with more skills like housing less (more) than those with fewer skills, the income elasticity estimate will be biased downwards (upwards). Our index also does not handle how earnings over the life-cycle may differ from permanent income.

from the average log wages households would earn in a typical city based on their human capital and other location-invariant characteristics.

The main testable restriction is that $\beta_1 + \beta_2 = 0$, which may be seen as a joint test of both demand theory and mobility.[16] When this restriction holds, the elasticity of substitution between housing and non-housing goods is $\sigma_D \equiv -(\hat{y}^j - \hat{x}^j)/(\hat{p}^j - \hat{c}^j) = -\epsilon^H_{y,p}/(1 - \bar{s_y})$, so that $\beta_1 = (1 - \bar{s_y})(1 - \sigma_D)$. When σ_D is less (greater) than one, housing demand is said to be price inelastic (elastic), and the expenditure share of housing rises (falls) with the relative price of housing, p/c. An advantage of the compensated specification is that it estimates the elasticity of substitution without reference to income, which our skill-index may not fully capture.

The general prediction for neutral quality-of-life amenities, with $\epsilon_{y,Q} = 0$, is that they are net substitutes for housing. However, when housing is a necessity, they increase the housing share as households take lower real incomes to live in more amenable areas: $\epsilon^H_{y,Q} - \epsilon_{m,Q} = -(1 - \epsilon_{y,m})\epsilon_{m,Q} < 0$. With an estimate of $\epsilon_{m,Q}$ — available from methods in Rosen (1979) or Albouy (2008) — the uncompensated effect of an amenity on housing demand is calculated by netting out this implied income effect using the formula $\epsilon_{y,Q} = \beta_4 + \beta_3 \epsilon_{m,Q}$.

3.4 Economies of Scale in Housing Consumption

Shared living quarters make housing consumption partly non-rival in nature. This "non-congestibility," proposed by Barten (1964), and explored by Deaton and Paxson (1998), can have potentially large price and income effects on housing demand. Here we sketch how to incorporate non-congestibility in the demand framework. The uncompensated demand function for housing *per household member* in a household of size n is:

$$y = n^{\phi} \tilde{y}\left(c, pn^{\phi-1}, \frac{m}{n}\right) \tag{6}$$

where $\tilde{y}$ is the housing demand function for a single-member household, and m/n is income per capita. The parameter ϕ governs the degree to which housing is congestible: $\phi = 1$ implies that

[16] If mobility does not hold, then the coefficients would not be of equal magnitudes. Income effects in the uncompensated elasticities would likely push coefficients on both housing and non-housing prices downwards.

housing is a purely congestible (or private) good. Log-linearizing this equation shows how the log expenditure share varies with prices, houshold size, and log income per capita, $\hat{m} - \hat{n}$:

$$\hat{s}_y = (\varepsilon_{y,p} + 1)\,\hat{p} + (\varepsilon_{y,w} - 1)\,(\hat{m} - \hat{n}) - (\varepsilon_{y,p} + \varepsilon_{y,m})\hat{c} - (1 - \phi)\,[\varepsilon_{y,p} + 1]\,\hat{n} \tag{7a}$$

$$= \alpha_1\,(\hat{p} - \hat{c}) + \alpha_3(\hat{m} - \hat{n} - \hat{c}) + \alpha_n \hat{n} \tag{7b}$$

It then follows that congestibility $\phi = 1 + \alpha_n/\alpha_1 \in [0, 1]$. Note this places a restriction on the estimates that $0 \geq \alpha_n \geq -\alpha_1$. In the appendix we demonstrate an analogous framework for compensated demand, with $\phi = 1 + \beta_n/\beta_1 \in [0, 1]$. The compensated framework is less appropriate for dealing with differences in household size given our use of the skill index in light of endogenous labor supply and household size choices.

3.5 Non-Homothetic Utility, Housing Share, and Cost-of-Living Functions

To allow the housing share to vary with both prices and income, we use the non-homothetic separable family CES (NH-CES) function from Sato (1977). Unlike other utility functions, it neatly separates out a substitution parameter σ, from a non-homotheticity parameter, γ, as well as a distribution parameter δ. We amend the function for imperfect household congestion, ϕ and neutral shifts in quality of life, Q:

$$U(x, y; n, Q) = Q \left[\frac{\delta \left(y n^{1-\phi} \right)^{\frac{\sigma-1}{\sigma}} + \theta_1}{\theta_2 - (1 - \delta) x^{\frac{\sigma-1}{\sigma}}} \right]^{\frac{\sigma}{\gamma(\sigma-1)}} \tag{8a}$$

where $\theta_1 = [1 - \sigma - \gamma\delta]/(\gamma\sigma)$ and $\theta_2 = [1 - \sigma - \gamma(\delta - 1)]/(\gamma\sigma)$. This function becomes a standard CES function (Arrow et al. 1961) as $\gamma \to 0$, and Cobb-Douglas (1928) if also $\sigma \to 1$. We show in the appendix that the housing share and cost-of-living index associated with this utility function are:

$$s_y(p,c,u;n,Q) = \frac{\delta^\sigma \left(\frac{p}{n^{1-\phi}}\right)^{1-\sigma}}{\delta^\sigma \left(\frac{p}{n^{1-\phi}}\right)^{1-\sigma} + (1-\delta)^\sigma c^{1-\sigma} \left(\frac{u}{Q}\right)^{\gamma(1-\sigma)}}. \tag{8b}$$

$$COL(p,c,u;n,Q) = \frac{u}{Q}\left[\delta^\sigma \left(\frac{p}{n^{1-\phi}}\right)^{1-\sigma} + (1-\delta)^\sigma c^{1-\sigma} \left(\frac{u}{Q}\right)^{\gamma(1-\sigma)}\right]^{\frac{1}{1-\sigma}} \tag{8c}$$

When $\gamma(1-\sigma) > 0$, households with higher utility consume less in housing, and need less income to compensate them for rises in p. Our restricted log-linear model maps to this utility function: when the right-hand side variables are demeaned, $\beta_0 = \sigma \ln \delta = \ln \bar{s}_y$, $\beta_1 = (1-\bar{s}_y)(1-\sigma)$, and $\beta_3 = -\gamma(1-\bar{s}_y)(1-\sigma)/\epsilon_{m,u}$, where $\epsilon_{m,u}$ is the expenditure elasticity with respect to u. The parameters are determined recursively with $\sigma = 1 - \beta_1/(1-e^{\beta_0})$, $\delta = e^{\beta_0/\sigma}$, $\gamma = -\epsilon_{m,u}\beta_3/\beta_1$.[17]

The cost-of-living index in (8c) requires that prices p and c are expressed in proportion to a reference level of prices $\bar{p}$ and $\bar{c}$. Furthermore, with a value of σ and a reference value for $\bar{s_y}$, e.g., the national average, the distribution parameter is set to $\delta = \{1 + [\bar{s_y}/(1-\bar{s_y})]^{(1/\sigma)}\}^{-1}$. We can incorporate a reference utility level for any household based on its housing consumption s_y^i and its quality of life, as 8b implies $(u/Q)^{\gamma(1-\sigma)} = \left[(1-s_y^i)\delta^\sigma \bar{p}^{1-\sigma}\right] / \left(s_y^i(1-\delta)^\sigma \bar{c}^{1-\sigma}\right)$. Local quality of life measures, Q, may be estimated using methods described in Albouy (2008).

3.6 Addressing Potential Biases in Elasticity Estimates

Several potential biases can arise in the estimation of the price and inccome elasticities of housing demand. Regarding the price elasticity, our approach corrects for potential biases that may arise from omitting non-housing prices, skill levels, and home-ownership. Additionally, the focus on cross-metro variation limits potential biases due to taste-based sorting.

Figure 5A shows that non-housing prices vary positively with housing prices. Suppose $\hat{c}^j = \rho\hat{p}^j + v^j$, where $\rho > 0$ and v^j is white noise. Substituting this projection into equation 5b, together with the elasticity of substitution, σ_D, gives that $\hat{\sigma}_D = 1 - \beta_1/[(1-e^{\beta_0})(1-\rho)]$. Thus, the higher

[17] Because the units of u and γ are not separately identified, we impose the restriction, $COL(1,1,u;n,Q) = 1$ to solve for γ and u/Q simultaneously.

is ρ, the more ignoring non-housing prices biases $\hat{\sigma}_D$ upwards.

Omitting the skill level of workers, ζ^j, can also bias estimates. As seen in figure 5B, higher-skilled households locate slightly more in high-rent areas.[18] If housing is a necessity, these households will exhibit a smaller housing share. Thus, higher-price areas have a low housing share relative to their rent level without this control, biasing estimated substitutability upward.

Another bias may stem from housing tenure selection. Suppose the propensity to rent rises with rent levels (possibly due to financing constraints), so that skilled households rent more in expensive cities. If housing is a necessity, and controls for skills (or utility) are incomplete, this could bias substitution elasticities upward. This suggests a control for the home-ownership rate.

Finally, there is the issue of unobserved taste-based sorting. Households that care more for housing should sort to areas where rents are low, negatively biasing the expenditure-rent gradient towards finding higher substitution. Such sorting behavior would likely cause the homogeneity restrictions to fail. To check, we compare estimates using rent variation across metros with those using variation within metros, where sorting is more likely. Nonetheless, this potential bias leads us to view our estimates of the elasticity of substitution as an upper bound on the true value.

The main concern in estimating the income elasticity of housing demand is attenuation bias from using current-period income. Indeed, our data include numerous observations for which rent paid exceeds reported income. Taking metro-level medians should greatly reduce biases produced by measurement error and transitory income. Nevertheless, cities themselves may be subject to transitory income shocks. The wage index measure is purged of any location effects and should suffer far less from these issues. It is still limited in that it only captures a snapshot of earnings over the life cycle.[19] Therefore, we view our preferred estimate of the income elasticity as a lower bound on the actual elasticity.

[18] Moretti (2013) finds a stronger skill-rent relationship using education only. Our measure includes race, experience, immigration status, and language ability.

[19] Classical measurement error in income implies that we observe $\hat{m}_*^j = \hat{m}^j + \eta^j$, where η^j is white noise. Defining $\lambda = 1 - var(\eta^j)/var(\hat{m}_*^j|\hat{p}^j, \hat{c}^j)$ as the reliability ratio, conditional on the other variables, the OLS estimate of α_3 will give $\lambda(\alpha_3 + 1) - 1$, and the inferred value of $\varepsilon_{y,m}$ is attenuated classically to zero by the factor λ. Haider and Solon (2006) estimate that as a measure of lifetime earnings, λ peaks in the middle of the life cycle at a value of about two-thirds.

3.7 The Housing Share, or "Affordability" as a Measure of Welfare

What housing shares, s_y, tell us about well-being or affordability hinges on the nature of the demand function. Housing must be a necessity for large shares to signal low well-being; it must be price inelastic for high shares to indicate high prices. But if households are mobile, high prices should reflect high wages or high quality of life. A household in an unsafe area with bad schools and long commutes may spend little on housing, but still be worse off than a household spending a larger share in an area with better amenities. If households are immobile, then high shares may indicate lower welfare, although differing amenities and tastes still complicate analysis.[20]

4 Data

The primary data source for our cross-sectional analysis is the 2000 Decennial Census microdata samples from IPUMS (Ruggles et al. 2004).[21] These data generate metro-level indices of income, m^j, predicted income, ζ^j, the rental-price index, p^j, and the housing share, s_y^j, as explained below. For the price of non-housing goods, we use a series from Carrillo et al. (2013), or "CEO," who construct the series from data by the American Chambers of Commerce Research Association (ACCRA).[22] In our time series analyses, we combine data from several sources. The 1970, 1980, 1990, and 2000 Decennial Censuses and 2001 through 2014 American Community Surveys provide data on the housing share, as do the 1984-2014 Consumer Expenditures Surveys (CEX) and personal consumption expenditures data from the Bureau of Economic Analysis for the years 1970 to 2014. Current Population Survey (CPS) data from 1970 to 2014 provide data on household size and composition, and Consumer Price Index data from the Bureau of Labor Statistics from the same years provide information on relative prices. The CPI for shelter from the BLS is

[20] Additionally, elderly households, particularly homeowners, may consume high amounts of housing because they have not adjusted from when their households were once larger.

[21] In the appendix we also consider the 1980 and 1990 Censuses and the combined 2007-2011 American Community Survey (ACS). Each represents 5 percent of the U.S. population. The metro-level indices are calculated for Primary Metropolitan Statistical Areas using 1999 Office of Management and Budget definitions. The Public-Use files are available for Public-Use Microdata Areas (PUMAs), are matched to metropolitan areas using standard geographic correlation techniques, which attempt to preserve the geography over different cross-sectional samples.

[22] These data begin in 1982, and so we use 1982 values for our 1980 specification.

based on observed rents, and rents imputed for owned units using a rental-equivalence approach.[23] The index is chain-weighted, accounting for changes in the geographic distribution of occupied houses. The BEA measure of housing expenditures imputes rental-equivalent measures for owner-occupied units. From the CEX, we take measures of average rental expenditures relative to all expenditures. Both datasets include owner-occupiers.[24]

4.1 Rental and Housing Expenditure Shares

We focus on median rental expenditure shares. Expenditures for owner-occupiers are complicated by difficulties in measuring user costs of housing and complications from savings. The rental share is the ratio of gross rents (including utility costs) to reported household income. Median shares circumvent aggregation issues and mitigate measurement problems such as under-reporting income, which can create very high shares for low income households.[25]

4.2 Price and Wage Indices

To calculate cross-sectional rental and house-price indices, we run regressions of the form $\ln(P^{ij}) = \alpha_P + \beta_P X^{ij} + \delta_P^j + \epsilon_P^{ij}$, where P^{ij} is the rent or imputed rent for unit i in area j. X_P^{ij} is a vector of housing-unit characteristics described in the appendix.[26] The estimated area indicators, δ_P^j, act as our inter-area housing price indices, p^j. To estimate a skill index, we run the wage regression

[23] The index is based primarily on a re-weighting procedure. The rental portion of the index may suffer from a downward bias, discussed in Crone et al. (2010).

[24] We use total expenditures as the denominator rather than income when we use the CEX, as it is closer to the ideal presented in the model.

[25] We also consider average and aggregate expenditure shares, equal to the sum of all rental payments divided by the sum of all tenant income. We consider two possible expenditure measures for owner-occupiers. The first is total monthly payments (or "cash-flow") related to housing, including mortgages, property taxes, and utilities. While this measure is appropriate for a static environment, it may diverge significantly from the true user-cost due to expected capital gains, mortgage terms, and net improvements relative to (unobserved) depreciation and maintenance costs. Most importantly, we do not observe income from home equity, which belongs on both the expenditure and income side of the equation. We also consider a measure of self-reported housing values relative to household income. Ideally, we would be able to model the decision to rent or own.

[26] We impute rents using the variable OWNCOST from the IPUMS microdata, which is the sum of mortgage and similar payments, real estate taxes, hazard insurance, utilities and fuels expenses, and condominium and mobile home fees where appropriate. When the regression includes both rented and owned units, X_P^{ij} includes tenure status interacted with every characteristic.

$\ln(W^{ij}) = \alpha_W + \beta_W X_W^{ij} + \delta_W^j + \epsilon_W^{ij}$, where W_{ij} is the hourly wage for person i in area j. X_{ij}^W is a vector of personal characteristics, and δ_j^W is a set of area fixed effects. The skill index ζ^j is from the corresponding moment (e.g., median or mean) of the $\hat{\beta}_W X_W^{ij}$.[27] The appendix covers additional indices for robustness.

5 Empirical Results

This section describes our estimates of the housing demand function. The heart of our estimation strategy, in tables 1 and 2, is to use cross-sectional variation across U.S. metropolitan areas. Table 3 displays the results from aggregate time series data.

5.1 Cross-sectional Evidence

5.1.1 Price and Income Elasticities: Estimates and Robustness

Table 1 presents metro-level estimates using the compensated model from equation (5b), using the log median rental share as the dependent variable in columns 1-4. Column 1 displays the results of a simple regression of the log median rental share on the rental price index, recovering a median expenditure share, s_y, of 22.5 percent, and an implied price elasticity, $\varepsilon_{y,p}$ of -0.83.[28]

Figures 6A and 6B illustrate the inter-metropolitan relationship between median expenditure shares and relative prices in the Census data. Figure 6A includes renters only, while 6B includes both renters and owners.[29] The regression line has slope $\beta_1 = -\beta_2$ in equation (5b), with $\beta_3 = \beta_4 = 0$ imposed. Both slopes are positive and statistically significant, indicating demand is price-inelastic: expenditure shares are higher in areas with more expensive housing. Figure 6B features a steeper slope and a tighter fit, although the housing share for home owners includes net investments and excludes implicit rents. The slope may be biased for more reasons than discussed above.

[27] Thus, raw wage differences across cities are the product of differences due to the area itself – compensating wage differentials for costs-of-living and amenities – and the local skills of the workforce, summarized by the wage index. Additional specifications use the average predicted values from the wage regressions.

[28] More technically, s_y is the geometric mean of the median expenditure share across metro areas.

[29] Accordingly, the former uses a price-index for rental units and the latter an index for all housing units.

As expected, the coefficient on rents, β_1, increases when the non-housing price index and the skill index are included in column 2 of table 1. The implied price and income elasticities are -0.7 and 0.7, respectively. A test of the homogeneity of demand does not reject the null hypothesis that $\beta_1 = -\beta_2$. Therefore, our preferred specification in column 3 imposes homogeneity of demand, recovering an elasticity of substitution, σ_D, of 0.69. The results in column 4 are largely unaffected by controlling for local home-ownership rates, suggesting that unobserved determinants of rentership do not bias our estimates.[30]

Column 5 uses the out-of-pocket expenditure share of home-owners as the dependent variable, while column 6 includes owners and renters, as in Figure 6B. The estimates in these columns imply smaller income and price elasticities than the results that are restricted to renters. The results in column 5 fail the homogeneity restriction, casting doubt on the reliability of the results that include homeowners, for whom it is more difficult to measure the theoretically appropriate concept of rents.[31] Column 7 uses the aggregate rental share, thereby weighting households in proportion to their expenditures. The results are similar to column 3, with slightly higher elasticities.

Column 8 presents results using within-metro variation at the PUMA level to examine household sorting. This specification suggests a lower income and higher price elasticity. The specification fails the homogeneity test, however, suggesting that either the non-housing cost data are not reliable within metros or that taste-sorting is biasing the estimates.

Overall, our preferred estimates in column 3 reveal an uncompensated price elasticity of roughly -2/3, an income elasticity near 2/3 and an elasticity of substitution near 2/3. Furthermore, the homogeneity restriction holds when using proper rental measures, even after controlling for home-ownership, suggesting our compensated demand framework based on household mobility is a useful estimation strategy.

[30]Davis and Ortalo-Magne's (2011) data for metro areas support an elasticity of substitution of 0.85. However, their index of rental costs differs from ours by controlling for commuting costs, and thus exaggerates the actual price differences faced by households (e.g. that suburban dwellers in the New York suburbs face Manhattan prices), biasing their results towards one. Their study does not account for income or non-housing prices.

[31]We have also tried imputing rents by adding utility to costs to a percentage of self-reported home values based on user costs. That percentage is either a uniform 6.2 percent, consistent with Albouy and Hanson (2014). The results are qualitatively similar, except that the specification in column 6 also fails the homogeneity test.

5.1.2 Household Size and Local Amenities

Table 2 incorporates household-size effects and local amenities both into the compensated demand specification in the previous section and into estimates of uncompensated regression equations modeled after equation (2a). The uncompensated regressions include a measure of household income per capita instead of the skill index in the compensated regressions, which proxies for earnings potential.

Columns 1 and 2 show the results of unrestricted Marshallian regression equations that include measures of household income per capita and household size as suggested by (7b), with an additional control for the fraction of household residents under the age of 18, following Deaton and Paxson (1998). The estimated price and income elasticities are somewhat smaller in column 1 than in the previous table. Although the specification in column 1 passes the homogeneity test easily, we remain suspicious of the estimated income elasticity, which is based on contemporaneous income, and therefore likely to be biased downward. The coefficient on log household size is negative: controlling for per-capita income, larger households consume less housing per capita. Combined with the price elasticity, the implied congestibility of housing is approximately 0.5, halfway between a pure public good and a pure private good within the household.[32]

Column 2 includes six commonly measured metropolitan amenities taken from Albouy (2008): mild winters, cool summers, sunshine, coastal proximity, hilliness, and clean air. Each is renormalized in standard deviations and signed so that a positive value is supposed to increase utility. None of the amenities is statistically significant at the 5-percent level.[33] The addition of so many amenities appears to have saturated the model to the point where the price and income estimates no longer satisfy homogeneity, which may be the result of imperfect measurement and the Iron Law of Econometrics (Hausman 2001).

Columns 3 and 4 consider compensated models that incorporate household size and composition effects. They replace the skill index with a skills-per-capita index. The coefficient on house-

[32]Intriguingly, this estimate is not far from equivalence scales that suggest using the square root of household size (OECD 2008), although here we only consider housing.

[33]Cool summers and coastal proximity are significant at the 10-percent level.

hold size is slightly smaller than in column 1 and less precise. Indeed, it is no longer statistically significantly different from zero, perhaps reflecting the greater difficulties of applying a Barten-style model in a compensated demand framework. The implied congestion parameter is small, but is not distinguishable from the earlier value of 0.5 at normal significance levels.

In column 4, we see stronger evidence that some amenities may impact housing demand. The typically positive coefficients provide indirect evidence that housing is a necessity, as higher amenities should reduce real money incomes holding utility constant. Indeed, the estimates in column 5 show the results of a regression of estimated quality of life, adapted from Albouy (2008), on the regressors from column 4.[34] The estimates confirm that all of these amenities increase the willingness-to-pay of households to live in a given area. A one standard deviation increase in each amenity appears to lower real income between 0.6 percent, in the case of clean air, and 2.5 percent, in the case of proximity to the coast.

Once the income effects of these amenities are netted out, as seen in column 6, their effects on housing demand are more striking. The effect of the average slope of the land, or "hilliness", is decidedly stronger than in column 2, suggesting that households may indeed demand larger houses in hilly areas after controlling for the higher prices there. This could be due to residents wanting to take advantage of better views from their homes, or enjoying the greater visibility of their homes. Furthermore, there is stronger evidence that housing demand in places with extreme heat also tends to be greater. This pattern may arise if heat induces residents to spend more time on their property, indoors or out. One caution with the amenity estimates is that they are based on a regression that no longer satisfies the homogeneity restriction, perhaps from multi-collinearity among regressors.

Although they stretch the limits of the data, we view the results of table 2 as broadly consistent with our preferred results in table 1. There is some evidence that housing is a partially public good within the household, with a congestibility parameter of about one half. There is also suggestive evidence that some amenities, like hilliness and extreme heat, increase demand for housing.

[34] This is based on a regression of a quality of life/willingness-to-pay index on the observed amenities. Following Albouy (2008), this index balances costs-of-living relative to after-tax nominal income gains given by $\hat{Q}^j = s_y \hat{p}^j + s_c \hat{c}^j - (1-\tau) s_w \hat{w}^j$, where τ is the marginal tax rate, s_w is the average income share from labor, and $\hat{w}^j$ is the estimated impact of location j on renters' wages.

5.2 Time Series Evidence

5.2.1 Time Series Estimates

Table 3 presents uncompensated demand estimates using the time series data presented in figures 1 and 3 for the years 1970-2014. These specifications use nominal prices and incomes, and thus (perhaps ambitiously) *estimate* an inflation index that balances housing versus non-housing goods from the observed behavior of households. We consider two different indices for the relative price of housing. The first uses the BLS's official measures, while the second incorporates the revisions proposed by Boskin (1996) and Crone, Nakamura, and Voith (2010).

We focus on restricted models satisfying homogeneity of degree zero in prices and income, which are not rejected by formal tests in our preferred specifications. The specifications include two additional terms. One is a linear time trend, t, that may capture secular changes in household preferences, for instance due to cohort effects, or increasing complementarity with local amenities as households have shifted locations.[35] The second term is the logarithm of household size, $\ln(n)$. The multi-collinearity between prices, incomes, and household size pushes the limit of what the time-series can identify.

Estimates from the BEA numbers with the official CPI, shown in column 1, imply an own-price elasticity of -0.55, an income elasticity of 0.61, and household size effect of -0.36. These are not far from the uncompensated estimates in table 2 and pass the homogeneity test. The estimated time trend is 0.007 per year, which we suspect may come from an under-estimate of the income elasticity. This result is consistent with the findings of Aguiar and Hurst (2013). The coefficient on log household size suggests a low degree of congestibility in housing demand.

The estimates in column 2, which use the revised relative price index, suggest larger price and income elasticities that are even closer to the cross-sectional estimates, but nonetheless still pass the homogeneity restriction. They also exhibit a similar estimated time trend, although the coefficient on log household size becomes positive, contrary to intuition and our previous results.

[35]The time trend may also reflect simple measurement error resulting from limitations in the data and its ability to identify low-frequency responses in housing consumption from shifting prices and income.

Columns 3 and 4 use the expenditure shares from the CEX for renters. These estimates are less precise and produce income and substitution elasticities much closer to one. Using the revised relative price index does not change the estimates substantially. Again, the estimated time trend is about 0.01 per year. The estimated coefficients on log household size are too imprecise to draw meaningful conclusions about the congesitibility of housing demand.

5.2.2 Explaining Changes in the Housing Share over Time

In table 4, we decompose the growing share of income spent on housing discussed in the introduction. Rearranging (4) and replacing Q with n and t, we have

$$\hat{s}_y = (1 - \bar{s_y} + \epsilon^H_{y,p})(\hat{p} - \hat{c}) + (\epsilon_{y,m} - 1)[(\hat{m} - \hat{c}) - \bar{s_y}(\hat{p} - \hat{c})] + \alpha_n \hat{n} + \alpha_t t + e \tag{9}$$

The first component represents the change due to the pure compensated price effect. This effect is positive when the relative price of housing increases if $\sigma < 1$, as $1 - s_y + \epsilon^H_{y,p} = (1 - s_y)(1 - \sigma)$. The second component is the income effect, from a parallel rise in the budget set, making the proper adjustment for changes in relative prices. The third component, $\alpha_n \hat{n}$, accounts for changes in household size; the fourth, $\alpha_t t$, the estimated time trend; and the fifth, e, is a residual.

Table 4 explains the overall increase in the log housing share of 6.7 percent (just under 2 percentage points) in the BEA numbers from 1970 to 2014. Column 1 uses parameters close to our preferred cross-sectional estimates in column 3 of table 1, with an elasticity of substitution and income elasticity of housing demand of two-thirds. The congestibility parameter ϕ is taken to be one-half, in line with the results in columns 1 and 2 of table 2. The specification incorporates the CPI numbers revised for both the CNV and Boskin critiques. The compensated price effect accounted for a nearly 23 percentage point increase in the housing share under this specification, while the income effect reduced it by by nearly 45 points. Shrinking household sizes also pushed up the housing share by 4 points. Nonetheless, the parameters from the cross-sectional estimates leave a large part of the increase in the housing share unexplained, as relatively easy substitution from housing to other goods leaves the compensated relative price effect outweighed by a large income effect.

Column 2 parametrizes the numbers by shading our estimates slightly in the direction where biases are most likely: thus $\sigma_D = 0.5 < 0.667$, and $\epsilon_{y,m} = 0.833 > 0.667$. ϕ is kept at 0.5. Column 2 also uses the CPI numbers revised for both the CNV and Boskin critiques. Under this specification, the relative price effect and the household size effect combine to explain an increase in the housing share of over 39 points. After the subtracting the income effect, the net effect is 17 points, slightly *over*-predicting the observed change of nearly 7 points, which is left in the time trend and residual. Thus, only modest changes to our cross-sectional parameter estimates are necessary to produce the observed increase in housing's share of expenditure.

Column 3 maintains the revised relative price index but uses the parameters estimated from the time series evidence in column 2 of table 3. The primary difference from the cross-sectionally estimated parameters is that the household congestion parameter is larger than one, contrary to what would be predicted by theory. Accordingly, the decrease in average household size led to an 8 point increase in the housing share in this specification, reducing the specification's ability to explain the increase in the housing share.

Column 4 uses our preferred parametrization from column 2 in combination with the official CPI. Using the official CPI implies both a smaller increase in the relative price of housing and slower growth in real incomes. The income effect's weaker tendency to decrease the housing share outweighs the price effect's weaker tendency to increase it, leading to a predicted change of approximately zero under these parameters. Column 5 uses the parameters from the time series estimates in column 1 of table 3, which also use the official CPI. The income effect predicts a very large fall in the housing share, which is offset only partially by a small price effect.

Finally, column 6 applies our preferred parametrization from column 2 along with the revised CPI to the CEX data from 1984 to 2014. The CEX data show a much larger increase in the housing share, of nearly 24 percent. Notably, the income effect is very small in the CEX data, as real incomes grew just 24 percent. Together, the relative price effect and slight decrease in average household size can account for just over half of the observed increase in the housing share in this specification.

Overall, the results in table 4 indicate that economic fundamentals are able to account for the observed increase in housing's share of expenditure since 1970 when Boskin's (1996) and Crone, Nakamura, and Voith's (2010) critiques of the official CPI are considered. Slight modifications to reflect the likely biases in our cross-sectional estimates of the housing demand function are sufficient to explain more than the entirety of the observed increase in the housing share.

6 Applying Estimates of Housing Demand

6.1 Changes in Housing Affordability, 1970 to 2010

Table 5 shows that since 1970, the percentage of households facing "extreme" (over 50 percent) housing affordability burdens rose from 16 to 28 percent, while the share facing "moderate" (over 30 percent) burdens rose from 30 to 53 percent. The median expenditure share devoted to rent rose from 20 to 31 percent in that time.

To explain this decline in affordability, we consider five separate trends in the economy from 1970 to 2010. First is the change in household composition and age structure. We consider this demographic change by dividing households into 36 categories household type and age categories, defined by the mean age of adults.[36] The 2010 sample is re-weighted so that these groups have the same proportion as in 1970.

The second trend is the increase in income inequality, which has pushed the income of renters down relative to average incomes. To assess this effect, we construct a counterfactual income distribution that assigns each household the income it would have earned if all incomes had increased proportionally between 1970 and 2010.[37] We denote household i's counterfactual income $\tilde{m}_i$. We

[36]We define the six household types as one adult without children, one adult with children, two adults without children, two adults with children, three or more adults without children, and three or more adults with children. We define the six age categories as 18-24, 25-34, 35-44, 45-54, 55-64, and 65+, where age is defined as the mean age of the adults.

[37]Formally, we calculate household incomes at each percentile, $k = 1, ..., 100$ for years $t = 1970, 2010$, m_t^k, as

multiply $\ln(\tilde{m}_i/m_i)$ by the income effect $\epsilon_{y,m} - 1$ to determine increasing income inequality's effect on household i's log income share devoted to housing $(\epsilon_{y,m} - 1)\ln(\tilde{m}_i/m_i)$.

Third, we consider increasing rental dispersion across metro areas, as in Moretti (2013). Increased dispersion will reduce affordability if rent increases happen disproportionately in areas with more renters. We assume that households' incomes are compensated for relative price increases, and thus calculate the compensated response $(\epsilon^H_{y,p}+1-\bar{s_y})[\ln(p^j_{1970}/p^j_{2010})-\ln(\bar{p}_{1970}/\bar{p}_{2010})]$.[38]

Fourth, following our previous analysis, we consider changes in average real incomes from 1970 to 2010. The income effect on housing demand is the change in average income, after accounting for the change in non-housing prices, times the income effect, $\epsilon_{y,m} - 1$. We consider both the official and revised CPI in these calculations.

Fifth and finally, we consider changes in the national average rent level. We calculate what affordability would have been if average rents had not increased from 1970 to 2010 using the uncompensated price elasticity of housing demand to account for the behavioral response through the formula $(\epsilon_{y,p} + 1)\ln(\bar{p}_{1970}/\bar{p}_{2010})$. Lower rents increase affordability provided that $\epsilon_{y,p} > -1$. We again consider the official and revised CPI in these calculations.

Table 5 accounts for these factors' contributions to the 22 and 12 percentage point increases in households facing moderate and extreme affordability burdens, respectively, as well as the 11 point increase in the median expenditure share among renters from 1970 to 2010. We consider two of the parametrizations from table 4. The first uses income and substitution elasticities of two-thirds each, close to our estimated values; the second uses an elasticity of substitution of 0.5 and an income elasticity of 0.833.

The change in household composition had small effects under both parametrizations, reducing the share of households facing an extreme burden by 0.8 percentage points but having a negligible effect on the other two measures. Widening income inequality had larger effects on moderate than extreme burdens, increasing them by 2.9 and 1.0 points, respectively, in the first parametriza-

well as mean incomes, $\bar{m}_t$. Based on each household's observed income m_i, the counterfactual income is $\tilde{m}_i = m_i[(\bar{m}_{2010}/\bar{m}_{1970})/(m^k_{2010}/m^k_{1970})]$.

[38]Using uncompensated regressions would create a lower response. We use contemporary population distributions to calculate relative price changes.

tion, and by half as much in the second. Changes in rent dispersion did disproportionately impact renters, although increasing dispersion explains less than a one point increase in affordability burdens in either parametrization.

The largest drivers of affordability burdens are increases in average incomes and average rents. The income effects are muted using the official CPI because real incomes changed little. The increase in average relative rents drives 5.8 and 3.2 point increases in moderate and extreme burdens under the first parametrization, and somewhat larger increases of 7.6 and 4.3 points under the second.

Using the revised CPI, the increase in average incomes has a much larger impact, as implied real incomes rose considerably faster than under the official CPI. Under the first parametrization, rising real incomes accounted for 6.1 and 0.2 point decreases in the affordability burdens, while under the second, the implied reductions in the affordability burdens were 3.0 and 1.8 percent. The change in average relative rents is much more pronounced using the revised CPI, explaining 13.3 and 5.5 point increases in the affordability burdens using the first parametrization, which increase to 16.3 and 9.3 points under the second.

Taken together, the five factors we consider explain more than 70 percent of the increase in housing affordability burdens from 1970 to 2010 using the revised CPI under our preferred parametrization.

6.2 Utility and Expenditure Functions

The estimates from the previous sections are sufficient to identify the utility and expenditure functions in section 3.5. For illustration, we round the parameters based off of estimates from column 3 of table 1, setting $\sigma = 2/3, \delta = 1/8, \gamma = 4/3$, and $\phi = 1/2$. Using these values in equations (8a) and (8c) yields the following utility and cost-of-living functions:

$$U_{NH-CES} = Q\left(\frac{27 - 14x^{-1/2}}{2y^{-1/2}n^{-1/4} + 3}\right)^{3/2}, COL_{NH-CES} = \frac{u}{Q}\left[\frac{1}{4}p^{1/3} + \frac{3}{4}c^{1/3}\left(\frac{u}{Q}\right)^{4/9}\right]^{3} \quad (10)$$

The units of x and y are median income shares for renters, with baseline values of $x = 0.78$ and $y = 0.22$. We hope these functions will have direct applicability to quantitative models in urban, macro, or public economics involving the housing sector.[39]

6.3 Cost-of-Living Indices over Space and Time

The estimated housing demand parameters may be used to calculate realistic cost-of-living indices (COLIs). We compare an index that assumes perfectly price- and income-inelastic housing demand (COL1), an index that assumes Cobb-Douglas preferences over housing and other goods (i.e., unitary price and income elasticities of housing demand; COL2), an index that relaxes the assumption of a unitary price elasticity but maintains the assumption of homotheticity (i.e., a CES demand function; COL3), and an index that allows for a non-homothetic CES demand function at the median household income and one half the median household income (Separable CES; COL4). Figure 7 plots these indices against the relative price of housing (p^j/c^j) for realistic variations over time and space. Panel A uses parameters close to those estimated in column 3 of table 1, while panel B uses the alternative parametrization in tables 4 and 5, which is meant to account for likely biases in the cross-sectional estimates.

Figure 7 shows how the fixed housing demand measure overstates differences in cost-of-living by ignoring households' ability to substitute between housing and other goods according to their relative prices, while the Cobb-Douglas preference measure understates these differences by assuming that substitution is easier than in reality. The separable (non-homothetic) CES COLI is steeper than those that fail to account for income effects. For poorer households, the other COLIs understate the burden of living in expensive areas, and overstate it in poorer areas. The correct index accounts for how high-rent cities are especially expensive for the poor. Of course, the regular

[39] For quantitative and analytical purposes, researchers might do well to use a more stylized LES utility and COL functions. Using a standard parametrization, set the share $\delta = 1/6$, and minimum consumption $\underline{x} = 0$ and, $\underline{y} = 0.1$.

$$U_{LES} = Q\left(yn^{1/2} - 0.1\right)^{1/6} x^{5/6}, COL_{LES} = 0.1p + 0.9\frac{p^{1/6}c^{5/6}}{n^{1/12}}\frac{u}{Q} \tag{11}$$

This corresponds to a typical housing share of 25 percent, with the minimum being 10 percent of median income. For a typical household the own-price elasticity is -2/3 and the income elasticity is 2/3.

CES function could be adapted to poorer households simply by changing its distribution parameter δ. The advantage of the non-homothetic CES function is that it offers a continuous mapping of cost-of-living for any income group, properly referenced to a given city at a given point in time.

6.4 Real Income Changes over Time and Growing Real Income Inequality

Using the ideal price index to deflate changes in income over time has different effects than using a fixed bundle index for two reasons. First, because housing is a necessity, the welfare of poorer households is reduced more by increases in the price of housing. Second, our index corrects for "substitution bias" automatically, mitigating welfare reduction for all groups. Table 6 deflates the nominal changes at the 10th, 50th, and 90th percentiles of the household income distribution from 1970 to 2010 using our ideal cost-of-living index and a comparable fixed-price index. Once again, we consider the two parametrizations from tables 4 and 5 along with the official and revised CPIs.

Nominal incomes of households at the 90th percentile rose 27 percentage more quickly than those at the 50th percentile and 25 percent more quickly than those at the 10th percentile. With homothetic preferences, the differences in real income growth of households at those percentiles would have tracked the differences in nominal income growth. Our ideal index (COL4), using the first parametrization, shows that with the official CPI, households at the 10th percentile saw real income growth of 10.9 percent, while those at the 90th percentile saw growth of 46.1 percent. This implies that real income inequality grew by 35.2 percent between these groups, or 3 percent more than standard numbers would show. Using the revised CPI, our ideal index implies that real incomes at the 90th percentile grew 9.7 percent faster than a fixed consumption bundle would imply, as the income effect pushed down richer households' share of expenditure devoted to housing. The results suggest that real income inequality grew by 45.7 percent from 1970 to 2010.

These differences are graphed in Figure 8 for the 50th and 90th percentiles. With the revised CPI we see incomes rising for all groups, but noticeably more for the 90th percentile. The figure shows that using the revised CPI implies an even larger increase in real income inequality than implied by the official CPI.

The smaller substitution possibilities implied by the second parametrization implies that all households experienced weaker real income growth than suggested by the first parametrization. However, the growth of real income inequality is also smaller under this parametrization.

7 Conclusion

Our econometric framework reveals that a spatial framework may be useful for estimating demand systems for goods, such as housing, whose prices vary considerably across locations. Both temporal and spatial estimates suggest that uncompensated own-price and income elasticities are close to two-thirds in absolute value. Thus, unit elasticities are better approximations than zero elasticities, although neither extreme can explain the observed variation in housing consumption across metro areas and over time.[40]

Rising rents appear to be the primary driver of the rising housing share in the national income accounts, and, to a lesser degree, the affordability crisis in rental markets. Increasing inequality as well as declining household sizes also appear to play a role, but there is some room for a secular rise in housing consumption that basic economic modeling cannot explain.

The estimated non-homothetic CES utility and cost-of-living functions we provide should be useful for realistic and tractable quantitative modeling in several economic fields.[41] They suggest substantial roles for substitution as well as for non-homotheticity. Indeed, we find that expensive cities are even more expensive for the poor, thereby exacerbating affordability problems. Moreover, nationally rising rents over time have increased real-income inequality considerably, even while spatial trends have not (Moretti, 2013). These findings highlight the idea that the "affordability crisis" in housing is deeply tied to the overall well-being of households, particularly at the bottom of the income distribution. Therefore, policies and regulations that raise rents by creating

[40] Taste-based sorting across space may bias our estimates towards finding greater price elasticity, but a large role for sorting seems incompatible the results of our specification tests. Moreover, those tests provide unique support for spatial estimates, as well as indirect evidence that household mobility is great enough to ensure compensating price and income differentials.

[41] Our elasticity of substitution estimates are consistent with the assumptions made by Albouy and Stuart (2014) and Rappaport (2008), who do not consider non-homotheticity.

artificial shortages in housing supply (Glaeser and Gyourko 2013, Albouy and Ehlrich 2016) may have particularly concerning distributional consequences.[42]

References

[1] Aguiar, Mark, and Erik Hurst. 2013. "Deconstructing Lifecycle Expenditure." *Journal of Political Economy* 121(3): 437-492.

[2] Albouy, David. 2008. "Are Big Cities Really Bad Places to Live? Improving Quality-of-Life Estimates across Cities." NBER Working Paper No. 14981.

[3] Albouy, David. (2016) "What Are Cities Worth? Land Rents, Local Productivity, and the Total Value of Amenities." *Review of Economics and Statistics.*

[4] Albouy, David, and Gabriel Ehrlich. 2012. "Housing Productivity and the Socail Cost of Land-Use Restrictions." NBER Working Paper No. 18110.

[5] Albouy, David, and Andrew Hanson. 2014. "Are Houses Too Big or in the Wrong Place? Taxes Benefits to Housing and Inefficiencies in Location and Consumption." *NBER Tax Policy and the Economy.*

[6] Albouy, David, and Lue Bert. 2015. "Driving to Opportunity: Local Rents, Wages, Commuting Costs and Sub-Metropolitan Quality of Life." *Journal of Urban Economics* 89: 74-92.

[7] Albouy, David, and Bryan Stuart. 2014. "Urban Population and Amenities: The Neoclassical Model of Location." NBER Working Paper No. 19919.

[8] Arrow, K.J., H.B. Chenery, B.S. Minhas, R.M. Solow (1961) "Capital-Labor Substitution and Economic Efficiency." *Review of Economics and Statistics*, 43(3): 225-250.

[9] Barten, Anton P. "Family Composition, Prices and Expenditure Patterns." in Peter E. Hart, Gordon Mills, and John K Whitaker, eds. *Econometric Analysis for National Economic Planning*. London: Butterworths, 1964.

[10] Baum-Snow, Nathaniel and Ronni Pavan. 2013. "Inequality and City Size." *Review of Economics and Statistics* 95(5): 1535-1548.

[11] Black, Dan, Gates Gary, Sanders Seth, and Taylor Lowell. 2002. "Why Do Gay Men Live in San Francisco ?" *Journal of Urban Economics* 51(1): 54-76

[12] Black, Daniel, Natalia Kolesnikova, and Lowell Taylor (2009)"Earnings Functions When Wages and Prices Vary by Location." *Journal of Labor Economics* 27(21).

[42] See Davis et al. 2016 for potential impacts of high rents on the educational impacts of the poor.

[13] Blundell, Richard, Panos Pashardes and Guglielmo Weber "What do we Learn About Consumer Demand Patterns from Micro Data?" *The American Economic Review* 83(3): 570-597.

[14] Carrillo, Paul, Dirk Early, and Edgar Olsen. 2013. "A Panel of Interarea Price Indices for All Areas in the United States 1982-2012" Available at SSRN: http://ssrn.com/abstract=2378028 or http://dx.doi.org/10.2139/ssrn.2378028.

[15] Cheshire, Paul, and Sheppard Stephen. 1998. "Estimating the Demand for Housing, Land, and Neighbourhood Characteristics." *Oxford Bulletin of Economics and Statistics.* Department of Economics, University of Oxford, vol. 60(3): 357-82, August.

[16] Crone, Theodore M., Leonard I. Nakamura and Richard Voith. 2010 "Rents Have Been Rising, Not Falling, in the Postwar Period." *Review of Economics and Statistics* 92: 628-641.

[17] Cobb, Charles W. and Paul H. Douglas (1928) "A Theory of Production." *American Economic Review* 18(1): 139-165.

[18] Davis, Morris A., Jesse Gregory, Daniel A. Hartley, and Kegon T.K. Tan. 2016. "Neighborhood Choices and Neighborhood Effects." mimeo, Rutgers University. July 26, 2016.

[19] Davis, Morris A., and Jonathan Heathcote. "The price and quantity of residential land in the United States." Journal of Monetary Economics 54.8 (2007): 2595-2620.

[20] Davis, Morris, and Francois Ortalo-Magne. 2011. "Household expenditures, wages, rents." *Review of Economic Dynamics*(14): 248-261.

[21] Deaton, Angus. 1986. "Demand Analysis" in Z. Grilliches and M.D. Intriligator *Handbook of Econometrics*, 3: 1767-1839.

[22] Deaton, Angus and Christina Paxson. 1998. "Economies of Scale, Household Size, and the Demand for Food." *Journal of Political Economy*. 106: 897-930.

[23] De Leeuw, Frank. 1971. "The Demand for Housing: A Review of Cross-section Evidence." *The Review of Economics and Statistics* 53 (1): 1-10.

[24] Desmet, Klaus and Esteban Rossi-Hansberg. 2013. "Urban Accounting and Welfare." *American Economic Review* 103(6): 2296-2327.

[25] Eckhout, Jan. 2004. "Gibrats law for all cities." *American Economic Review* 94 (5): 1429-1451.

[26] Ermisch, J., J. Findlay, and K. Gibb, 1996, "The Price Elasticity of Housing Demand in Britain: Issues of Sample Selection. *Journal of Housing Economics* 5(1): 64-86.

[27] Engel, E. 1857. "Die Produktions- und Consumtionsverhaltnisse des Konigreichs Sachsen." reprinted with Engel (1895), Anlage I, 1-54.

[28] Guerreri, Veronica, Daniel Hartley, and Erik Hurst. 2013. "Endogenous Gentrification and Housing Price Dynamics." *Journal of Public Economics* 100: 45-60.

[29] Glaeser, Edward and Joseph Gyorko (2008) *Rethinking Federal Housing Policy: How to Make Housing Plentiful and Affordable.* Washington, DC: AEI Press.

[30] Glaeser, Edward L and Joseph Gyourko (2003). “The Impact of Building Restrictions on Housing Affordability.” *Federal Reserve Bank of New York Economic Policy Review*, 9, pp. 21-29.

[31] Goodman, Allen C. 1988. “An Econometric Model of Housing Price, Permanent Income, Tenure Choice, and Housing Demand.” *Journal of Urban Economics* 23 (May): 327353.

[32] Goodman, Allen C. 2002. “Estimating Equilibrium Housing Demand for “Stayers” of Housing Price, Permanent Income, Tenure Choice, and Housing Demand.” *Journal of Urban Economics*51: 1-24.

[33] Goodman, Allen C, and Kawai Masahiro. 1986. “Functional Form, Sample Selection, and Housing Demand.” *Journal of Urban Economics* 20: 155-167.

[34] Gyourko, Joseph, Christopher Mayer, and Todd Sinai, 2013. ”Superstar Cities,” *American Economic Journal: Economic Policy* 5(4): 167-99.

[35] Handbury, Jessie (2013) “Are Poor Cities Cheap for Everyone? Non-homotheticity and the Cost-of-Living across U.S. Cities.” mimeo.

[36] Hansen, Julia L. and Formby John P. 1998. “Estimating the Income Elasticity of emand for Housing:A Comparison of Traditional and Lorenz-Concentration Curve Methodologies.” *Journal of Housing Economics* 7: 328-342.

[37] Hanushek, Eric, and Quigley John. 1980. “The Dynamics of housing market: a stock adjustment model of housing consumption.” *Journal of Urban Economics* 6(1), pp. 90-111.

[38] Haider, Stephen and Gary Solon (2006) “Life-Cycle Variation in the Association between Current and Lifetime Earnings” *American Economic Review* 96: 1308-1320.

[39] Harmon, Oskar R. 1988. “The Income Elasticity of Demand for Single-Family Owner-Occupied Housing: An Empirical Reconciliation.” *Journal of Urban Economics* 24: 173-185.

[40] Hausman, Jerry. 2001. “Mismeasured Variables in Econometric Analysis: Problems from the Right and Problems from the Left.” *Journal of Economic Perspectives* 15(4): 5767.

[41] HUD PD & R Edge. “Secretary Donovan Highlights Convening on State of Americas Rental Housing.” January 2014. `http://www.huduser.org/portal/pdredge/pdr_edge_featd_article_012714.html`.

[42] Ioannides, Yannis M., and Jeffrey E. Zabel. 2003. “Neighborhood Effects and Housing Demand.” *Journal of Applied Econometrics* 18 (Sep): 563-584.

[43] Joint Center for Housing Studies of Harvard University. 2013. “America’s Rental Housing: Evolving Markets and Needs.” Mimeograph. Harvard University.

[44] Kau, James B., and Sirman C. F. 1979. “Urban Land Value Functions and the Price Elasticity of Demand for Housing.” *Journal of Urban Economics* 6(4): 112-121.

[45] La Cava, Gianni. 2016. “Housing Prices, Mortgage Interest Rates and the Rising Share of Capital Income in the United States.” Research Discussion Paper 2016-04, Reserve Bank of Australia.

[46] Malpezzi, Stephen, Gregory H. Chun, and Richard K. Green. “New PlacetoPlace Housing Price Indexes for US Metropolitan Areas, and Their Determinants.” Real Estate Economics 26.2 (1998): 235-274.

[47] Mayo, Stephen K. 1981. “Theory and Estimation in the Economics of Housing Demand.” *Journal of Urban Economics* 10: 95-116.

[48] Michaels, Guy, Ferdinand Raurch, and Stephen J. Redding. 2013. ”Urbanization and Structural Transformation.” *The Quarterly Journal of Economics* 127(2): 535-586.

[49] Mills E.S. 1967. “An aggregative model of resource allocation in a metropolitan area.” *American Economic Review* 57: 197-210.

[50] Moretti, Enrico 2013 ”Real Wage Inequality” *American Economic Journal: Applied Economics* 5(1): 65-103.

[51] Muth, Richard F. 1960. “The Demand for Non-Farm Housing.” in *The Demand for Durable Goods*, edited by Arnold C. Harberger, 29-96. Chicago, IL: University of Chicago Press, 1960. (Ph.D. dissertation)

[52] Piketty, Thomas. 2014. *Capital in the Twenty-First Century.* Cambridge, MA: Harvard University Press.

[53] Pollinsky, A. Mitchell, and Ellwood David T. 1979. “An Empirical Reconciliation of Micro and Grouped Estimates of the Demand for Housing.” *The Review of Economics and Statistics* 61(2): 199-205

[54] Rappaport, Jordan. 2008a. “A Productivity Model of City Crowdedness.” *Journal of Urban Economics* 65: 715-722.

[55] Reid, Margaret. 1962. *Housing and Income.* Chicago: University of Chicago Pres.

[56] Roback, Jennifer. 1982. “Wages, Rents, and the Quality of Life.” *Journal of Political Economy* 90: 1257-1278.

[57] Rognlie, Matthew. 2015. “Deciphering the Fall and Rise in the Net Capital Share: Accumulation or Scarcity?.” *Brookings Papers on Economic Activity* 2015(1): 1-69.

[58] Rosen, Harvey. 1985. “Housing Subsidies: Effects on Housing Decisions, Efciency, and Equity. In *Handbook of Public Economics*, vol. 1, edited by M. Feldstein and A. Auerbach. Amsterdam: North-Holland.

[59] Rosen, Sherwin. 1979. “Wages-based Indexes of Urban Quality of Life.” In *Current Issues in Urban Economics*, edited by P. Mieszkowski and M. Straszheim. Baltimore: John Hopkins Univ. Press.

[60] Ruggles, Steven, Matthew Sobek, Trent Alexander, Catherine A. Fitch, Ronald Goeken, Patricia Kelly Hall, Miriam King, and Chad Ronnander. 2004. *Integrated Public Use Microdata Series: Version 3.0*. Minneapolis: Minnesota Population Center.

[61] Saiz, Albert. 2010. ”The geographic determinants of housing supply.” *The Quarterly Journal of Economics* 125.3: 1253-1296.

[62] Sato, Ruozo (1977) “Homothetic and Non-Homothetic CES Production Functions” *American Economic Review* 67(4): 559-569.

[63] Schwabe, Hermann. 1868. ”Das Verhaltniss von Miethe und Einkom-men in Berlin.” in *Berlin und seine Entwickelung fur 1868* (Berlin, 1868): 264-67.

[64] Stigler, George J. 1954. “The Early History of Empirical Studies of Consumer Behavior.” *Journal of Political Economy* 62(2): 95-113

[65] Wright, Carroll. 1875. “Sixth Annual Report of the Bureau of Labor Statistics.”(Boston)

[66] Zivin, Joshua G., and Matthew Neidell, 2014. ”Temperature and the Allocation of Time: Implications for Climate Change.” *Journal of Labor Economics* 32(1): 1-26.

TABLE 1: COMPENSATED DEMAND FUNCTION ESTIMATES AT THE METROPOLITAN LEVEL USING 2000 CENSUS DATA

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Dependent Variable:	Log Median Rental Share	Log Median Rental Share	Log Median Rental Share	Log Median Rental Share	Log Median Spend. Share	Log Median Hous. Share	Log Aggreg. Rental Share	Log Median Rental Share
Panel A: Regression Results								
Rental/Housing Price Index	0.172	0.233	0.239	0.246	0.392	0.407	0.233	0.163
	(0.026)	(0.029)	(0.027)	(0.027)	(0.053)	(0.032)	(0.025)	(0.015)
Non-Housing Price Index		-0.185	-0.239	-0.246	-0.123	-0.407	-0.233	0.475
		(0.087)	(0.027)	(0.027)	(0.157)	(0.032)	(0.025)	(0.285)
Skill/Predicted Wage Index		-0.298	-0.282	-0.287	-0.511	-0.511	-0.173	-0.529
		(0.108)	(0.112)	(0.112)	(0.148)	(0.132)	(0.115)	(0.027)
Homeownership Rate				0.079				
				(0.061)				
Constant	-1.492	-1.492	-1.492	-1.492	-1.703	-1.628	-1.640	-1.495
	(0.006)	(0.004)	(0.004)	(0.004)	(0.008)	(0.006)	(0.005)	(0.001)
Sample Size (number of areas)	331	331	331	331	331	331	331	1655
Adjusted R-squared	0.297	0.518	0.519	0.524	0.577	0.674	0.432	0.109
Homogeneity of Demand Restricted	No	No	Yes	Yes	No	Yes	Yes	No
Unconstrained Sum of Housing and Non-Housing Price Index Coefficients		0.047	0.047	0.043	0.269	0.127	0.065	0.622
		(0.071)	(0.071)	(0.066)	(0.127)	(0.102)	(0.082)	(0.287)
P-value of Test of Homog. of Demand		0.504	0.504	0.519	0.035	0.212	0.429	0.031
Sample	Renters Only	Renters Only	Renters Only	Renters Only	Owners Only	Renters and Owners	Renters Only	Renters Only
Unit of Observation	Metro	Metro	Metro	Metro	Metro	Metro	Metro	PUMA
Panel B: Implied Demand Paramters								
Geometric Mean Expenditure Share	0.225	0.225	0.225	0.225	0.182	0.196	0.194	0.224
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.000)
Uncompensated Own Price Elasticity of Housing Demand	-0.828	-0.700	-0.697	-0.689	-0.515	-0.492	-0.734	-0.719
	(0.026)	(0.043)	(0.042)	(0.043)	(0.073)	(0.052)	(0.037)	(0.019)
Income Elasticity of Housing Demand	1.000	0.702	0.718	0.713	0.489	0.489	0.827	0.471
	Restricted	(0.108)	(0.112)	(0.112)	(0.148)	(0.132)	(0.115)	(0.027)
Elasticity of Substitution Between Housing and Consumption Goods σ			0.691	0.682		0.493	0.711	0.790
			(0.035)	(0.035)		(0.040)	(0.032)	(0.019)
Distribution Parameter δ			0.116	0.112		0.037	0.100	0.151
			(0.012)	(0.012)		(0.010)	(0.010)	(0.007)
Non-homotheticity Parameter γ			1.178	1.165		1.254	0.743	3.252
			(0.444)	(0.430)		(0.272)	(0.484)	(0.241)

Robust standard errors reported in parentheses. The predicted wage index is based on the wage level predicted by education, experience, race, immigrant status, occupation, and industry, partialing out the effect of location. Homogeneity of demand test is for whether the coefficients on the rental/housing price index and the non-housing price index sum to zero. All regressions include controls for the log trimean of household size. All regressions exclude households that include college students.

TABLE 2: UNCOMPENSATED AND COMPENSATED DEMAND FUNCTIONS WITH HOUSEHOLD SIZE AND AMENITY SHIFTERS - 2000 CENSUS DATA, RENTERS ONLY

	Uncomp-ensated	Uncomp-ensated	Compensated	Compensated	Amenity Value	Housing Effects
	(1)	(2)	(3)	(4)	(5)	(6)
Dependent Variable:	Rental Share	Rental Share	Rental Share	Rental Share	Qual. of Life	Adj. Rental Share
Panel A: Regression Results						
Rental/Housing Price Index	0.553	0.529	0.233	0.149		
	(0.023)	(0.033)	(0.029)	(0.025)		
Non-Housing Price Index	-0.037	0.105	-0.185	0.143		
	(0.064)	(0.060)	(0.087)	(0.084)		
Log Trimean Household Income per Capita	-0.533	-0.501				
	(0.036)	(0.036)				
Household Skill Index			-0.298	-0.300		
			(0.108)	(0.093)		
Log Trimean Household size	-0.299	-0.281	-0.204	-0.312		
	(0.045)	(0.039)	(0.147)	(0.117)		
Trimean Fraction of Children	0.047	0.058	0.291	0.182		
	(0.122)	(0.096)	(0.177)	(0.129)		
Mild Winters (Minus Heating Degree Days in 1000s)		0.002		0.011	0.007	0.009
		(0.005)		(0.008)	(0.007)	(0.005)
Cool Summers (Minus Cooling Degree Days in 1000s)		-0.010		-0.016	0.017	-0.021
		(0.006)		(0.008)	(0.006)	(0.005)
Percent of Annual Sunshine Possible		-0.004		0.012	0.018	0.006
		(0.005)		(0.006)	(0.004)	(0.005)
Proximity to Coast (Log Inverse Distance)		-0.007		0.002	0.025	-0.005
		(0.004)		(0.006)	(0.004)	(0.005)
Hilliness (Average Slope of Land)		0.005		0.019	0.009	0.016
		(0.003)		(0.004)	(0.003)	(0.003)
Clean Air (minus median AQI)		0.003		0.007	0.006	0.005
		(0.002)		(0.003)	(0.003)	(0.003)
Constant	-1.492	-1.492	-1.492	-1.492	-0.018	
	(0.003)	(0.002)	(0.004)	(0.003)	(0.004)	
Sample Size (number of areas)	331	329	331	329	329	
Adjusted R-squared	0.807	0.817	0.518	0.655	0.439	
Unconstrained Sum of Price and Income Coefficients	-0.017	0.133	0.047	0.292		
	(0.053)	(0.055)	(0.071)	(0.078)		
P-value of Test of Homog. of Demand	0.744	0.016	0.504	0.0002		
Panel B: Implied Demand Parameters						
Geometric Mean Expenditure Share	0.225	0.225	0.225	0.225		
	(0.001)	(0.001)	(0.001)	(0.001)		
Uncompensated Own Price Elasticity of Housing Demand	-0.447	-0.471	-0.700	-0.784		
	(0.023)	(0.033)	(0.043)	(0.038)		
Income Elasticity of Housing Demand	0.467	0.499	0.702	0.700		
	(0.036)	(0.036)	(0.108)	(0.093)		
Congestion of Housing Consumption φ	0.458	0.468	0.121	-1.095		
	(0.078)	(0.073)	(0.584)	(0.718)		

All specifications include renters only. Robust standard errors in parentheses. Test of homogeneity of demand for the uncompensated regressions is that the coefficients on both price indices and income sum to zero. Rental share in column 6 is adjusted to reflect quality-of-life's effects on real income as discussed in section 5.1.2 and footnote 33.

TABLE 3: NATIONAL HOUSING DEMAND OVER TIME - ESTIMATES

Data Source:	BEA	BEA	CEX	CEX
Consumer Price Index:	Official	Revised	Official	Revised
Dependent Variable:	Log Aggregate Housing Expenditure Share			
	(1)	(2)	(3)	(4)
Panel A: Restricted Regression Results				
Log CPI-U: Shelter minus Log CPI-U: All Items Less Shelter	0.452	0.351	0.171	0.177
	(0.042)	(0.034)	(0.140)	(0.131)
Log Average Income/Expenditures Per Capita minus Log CPI-U: All Items Less Shelter	-0.389	-0.324	-0.205	-0.207
	(0.060)	(0.051)	(0.173)	(0.172)
Linear Time Trend (years)	0.007	0.008	0.009	0.010
	(0.002)	(0.002)	(0.001)	(0.002)
Log Household Size	-0.359	0.333	0.760	0.819
	(0.093)	(0.125)	(0.698)	(0.683)
Constant	-1.717	-1.717	-1.222	-1.222
	(0.003)	(0.003)	(0.004)	(0.004)
Sample size (years)	45	45	31	31
Sample Period	1970-2014		1984-2014	
P-value of Test of Homogeneity of Demand	0.455	0.515	0.857	0.743
Panel B: Implied Demand Parameters				
Geometric Mean Expenditure Share	0.180	0.180	0.295	0.295
	(0.001)	(0.001)	(0.001)	(0.001)
Uncompensated Own-Price Elasticity of Housing Demand	-0.548	-0.649	-0.829	-0.823
	(0.042)	(0.034)	(0.140)	(0.131)
Income Elasticity of Housing Demand	0.611	0.676	0.795	0.794
	(0.060)	(0.052)	(0.173)	(0.172)

Newey-West standard errors reported in parentheses. Income/expenditure measure in per capita terms. Homogeneity of demand requires that the coefficients on log CPI-U for shelter, log CPI-U for all items less shelter, and log real household income sum to zero. The restricted regressions shown impose this constraint making one estimate redundant. For non-BEA series, a moving average with weight of 0.5 for the year after and the year before is used. Columns 2 and 4 revise the CPI-U for shelter according to Crone, Nakamura, and Voith (2010), and also revise the CPI-U for Shelter and All Items Less Shelter for the biases reported in the Boskin Commission (1996).

TABLE 4: DECOMPOSING INCREASES IN THE HOUSING SHARE OVER TIME

	(1)	(2)	(3)	(4)	(5)	(6)
Data Source	BEA	BEA	BEA	BEA	BEA	CEX
Consumer Price Index	Revised CPI	Revised CPI	Revised CPI	Official CPI	Official CPI	Revised CPI
Parameters Used:	Cross-Sectional Estimates	Low Subst./Weak Income Eff.	Time Series Estimates	Low Subst./Weak Income Eff.	Time Series Estimates	Low Subst./Weak Income Eff.
Elasticity of Substitution	0.667	0.500	0.643	0.500	0.534	0.500
Income Elasticity	0.667	0.833	0.676	0.833	0.611	0.833
Household Congestion	0.500	0.500	1.949	0.500	0.206	0.500
Observed Change:						
Change in Log Relative Prices	0.828	0.828	0.828	0.306	0.306	0.351
Change in Log Deflated Incomes	1.339	1.339	1.339	1.128	1.128	0.237
Change in Log Household Size	-0.239	-0.239	-0.239	-0.239	-0.239	-0.075
Change in Housing Share	0.067	0.067	0.067	0.067	0.067	0.239
Sample Years	1970-2014	1970-2014	1970-2014	1970-2014	1970-2014	1984-2014
Change in Housing Share Attributable to:						
Compensated Relative Price Effect	0.226	0.340	0.242	0.125	0.117	0.124
Income Effect	-0.446	-0.223	-0.434	-0.188	-0.439	-0.039
Household Size Effect	0.040	0.053	-0.080	0.053	0.086	0.015
Time Trend	0.272	-0.087	0.341	0.065	0.310	0.137
Residual	-0.025	-0.015	-0.002	0.012	-0.008	0.002
Total Unexplained	0.247	-0.102	0.339	0.077	0.303	0.140

Parameters in Estimate 1 (column 5) taken from Table 3, column 1. Parameters in Estimate 2 (column 3) taken from Table 3, column 2. Parameterization 1 allows for moderate substitutability between housing and other goods and assumes housing is a necessary good. Parameterization 2 assumes weaker substitutability and income effects.

TABLE 5: UNDERSTANDING INCREASES IN HOUSING AFFORDABILITY BURDENS FOR RENTERS, 1970-2010

	Median Exependiture Share (1)	Share with Moderate Burden (over 30%) (2)	Share with Extreme Burden (over 50%) (3)
Renter Households in 2010	0.306	0.524	0.276
Parameterization 1: Income Elasticity 2/3, Elasticity of Substitution 2/3			
1. Undoing Changes in Household Composition	0.306	0.522	0.268
2. Undoing Increases in Income Inequality	0.293	0.493	0.258
3. Undoing Changes in Relative Rents	0.291	0.489	0.254
4A. Undoing Increase in Average Income (Revised CPI)	0.321	0.550	0.257
5A. Undoing Increase in Average Rents (Revised CPI)	0.256	0.416	0.202
4B. Undoing Increase in Average Income (Official CPI)	0.292	0.492	0.257
5B. Undoing Increase in Average Rents (Official CPI)	0.264	0.434	0.224
Parameterization 2: Income Elasticity 5/6, Elasticity of Substitution 1/2			
1. Undoing Changes in Household Composition	0.306	0.522	0.268
2. Undoing Increases in Income Inequality	0.299	0.507	0.263
3. Undoing Changes in Relative Rents	0.296	0.500	0.258
4A. Undoing Increase in Average Income (Revised CPI)	0.311	0.530	0.276
5A. Undoing Increase in Average Rents (Revised CPI)	0.232	0.367	0.183
4B. Undoing Increase in Average Income (Official CPI)	0.297	0.502	0.259
5B. Undoing Increase in Average Rents (Official CPI)	0.261	0.426	0.216
Renter Households in 1970	0.197	0.303	0.156

Notes: Datasets used are 1970 Census and pooled 2009-2011 American Community Survey. Moderate burden is defined as an expenditure share on housing in excess of 30%; extreme burden is defined as expenditure share in excess of 50%. Counterfactual 1 assumes no change in household composition 1970-2010. Counterfactual 2 assumes no increase in income inequality 1970-2010. Counterfactual 3 additionally assumes no increase in dispersion of rents across metro areas 1970-2010. Counterfactual 4 additionally assumes no increases in average incomes 1970-2010, deflated by CPI. Counterfactual 5 assumes no increase in average rents 1970-2010, deflated by CPI.

TABLE 6: CHANGES IN REAL INCOMES 1970-2010 BY INCOME PERCENTILES IDEALLY DEFLATED

Household Position	Income Ratio 2009/1970	Ideal Deflator	Ideal Deflated Income	Deflated Fixed Bundle	Ideal Correction to Fixed
Parameterization 1: Income Elasticity 2/3, Elasticity of Substitution 2/3					
Panel 1A: Official CPI					
10th Percentile	6.103	5.504	1.109	1.120	-0.011
50th Percentile	6.002	5.440	1.103	1.102	0.002
90th Percentile	7.869	5.386	1.461	1.444	0.017
Panel 1B: Revised CPI					
10th Percentile	6.103	4.862	1.255	1.253	0.002
50th Percentile	6.002	4.717	1.273	1.232	0.040
90th Percentile	7.869	4.596	1.712	1.615	0.097
Parameterization 2: Income Elasticity 5/6, Elasticity of Substitution 1/2					
Panel 2A: Official CPI					
10th Percentile	6.103	5.473	1.115	1.120	-0.005
50th Percentile	6.002	5.442	1.103	1.102	0.001
90th Percentile	7.869	5.414	1.453	1.444	0.009
Panel 2B: Revised CPI					
10th Percentile	6.103	4.825	1.265	1.253	0.012
50th Percentile	6.002	4.753	1.263	1.232	0.031
90th Percentile	7.869	4.688	1.678	1.615	0.063

Income ratio in nominal terms from Census data. Ideal deflator uses estimated COL4 index and fixed-bundle deflator uses COL1 index as described in section 6.3. Revised CPI uses Boskin and CNV revisions described in the text. Ideal correction takes difference.

Figure 1A: Personal Consumption Expenditures by Major Category, 1959-2014

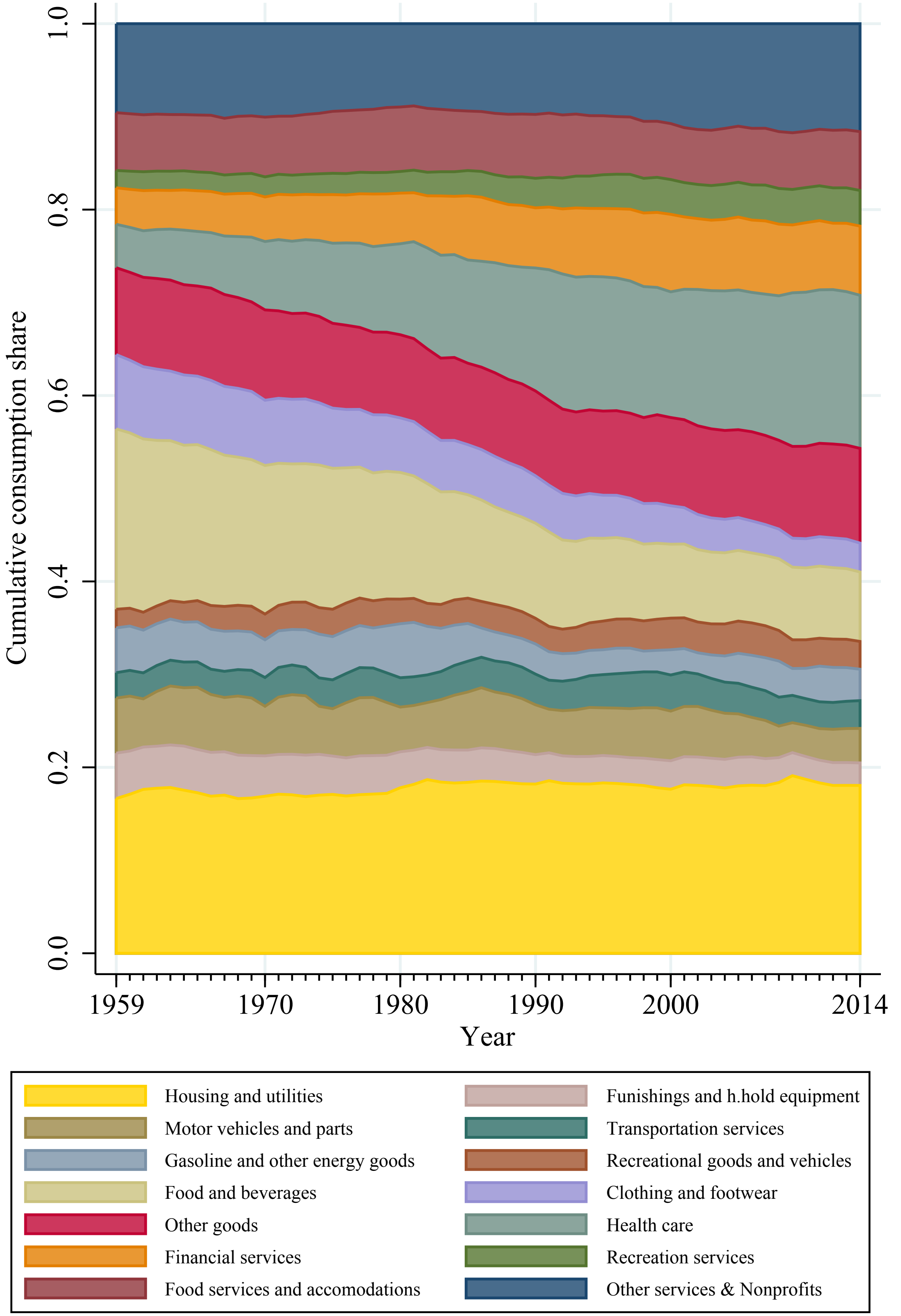


Figure 1B: Expenditure Share on Housing 1960-2015

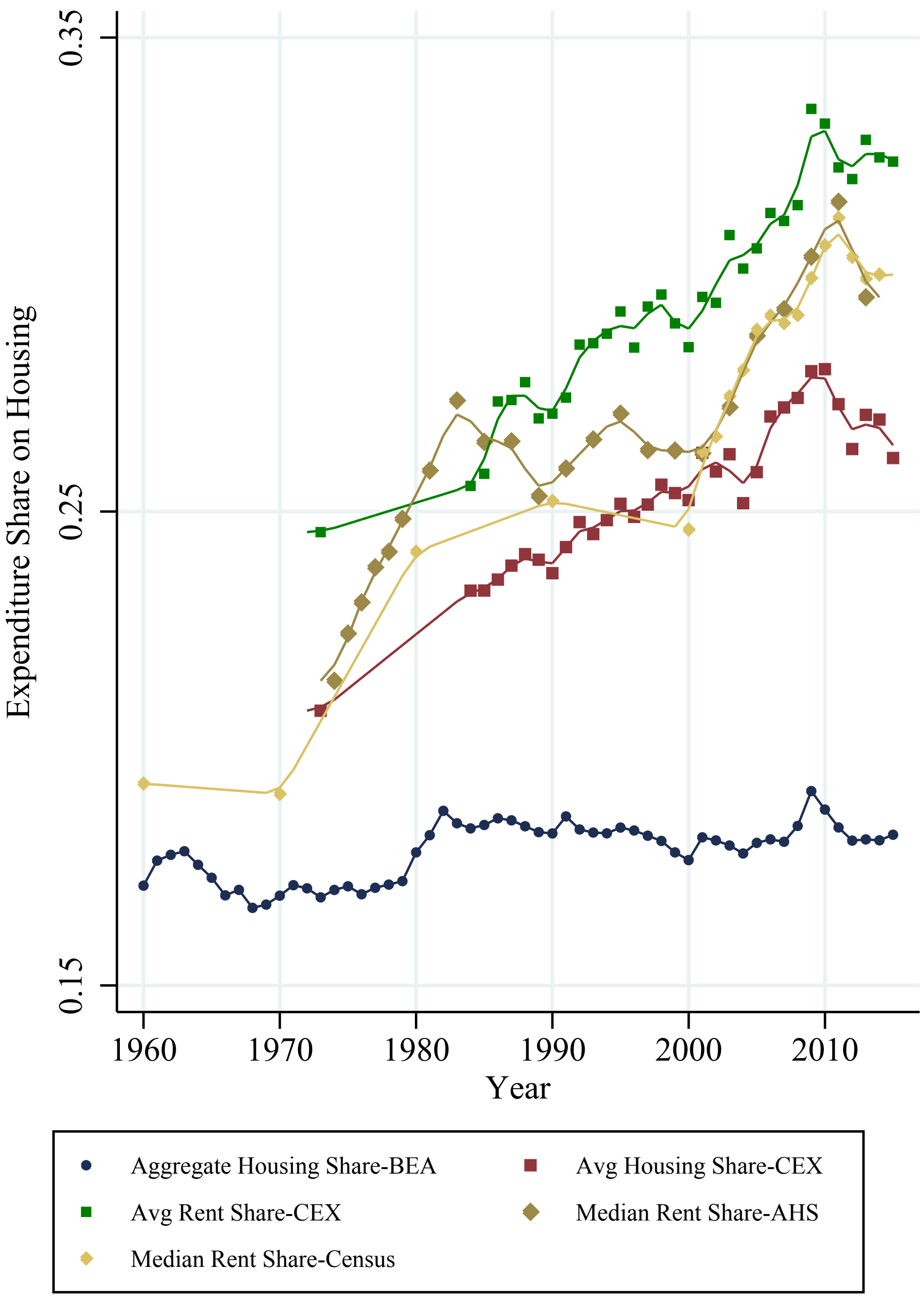
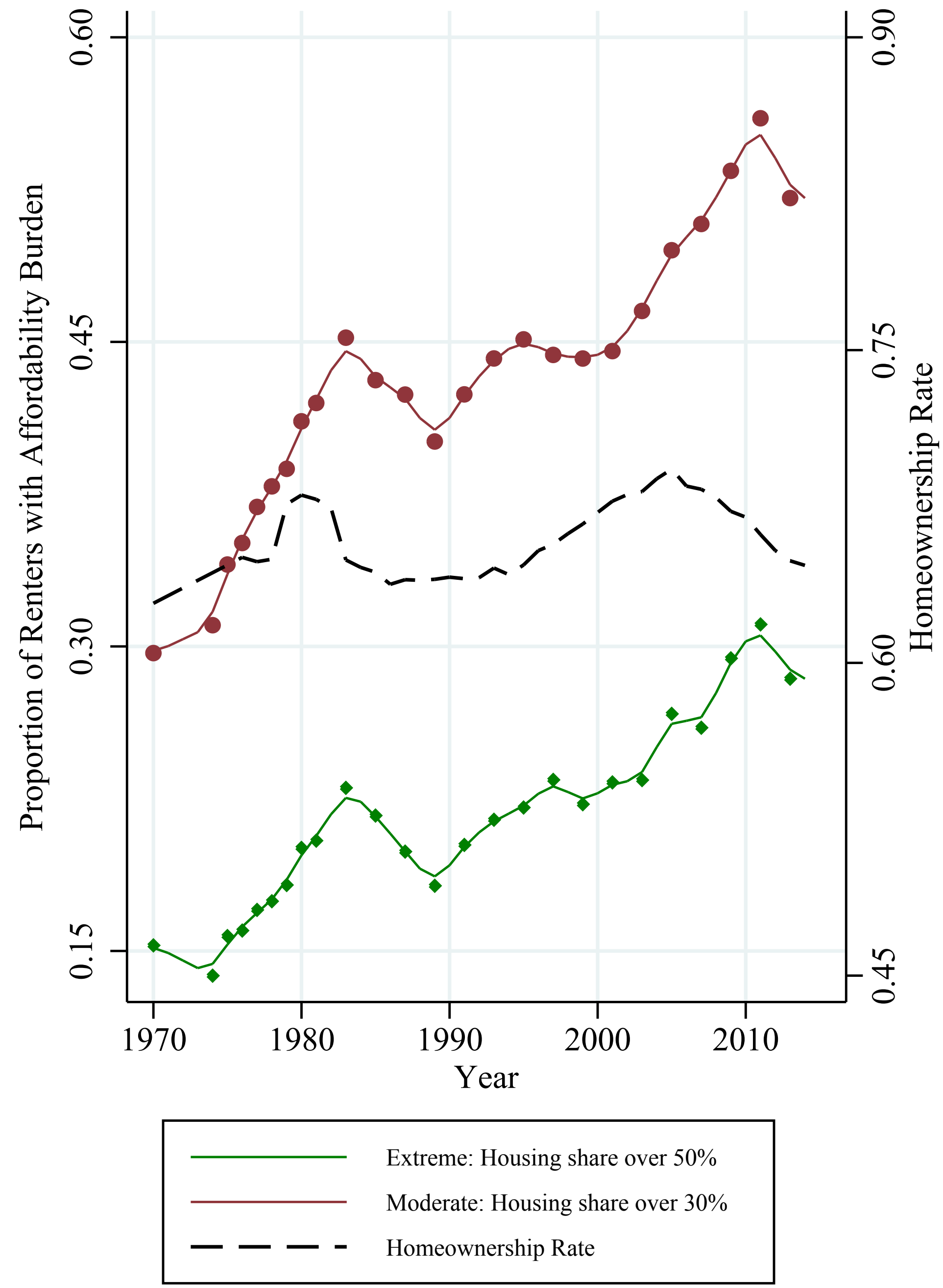
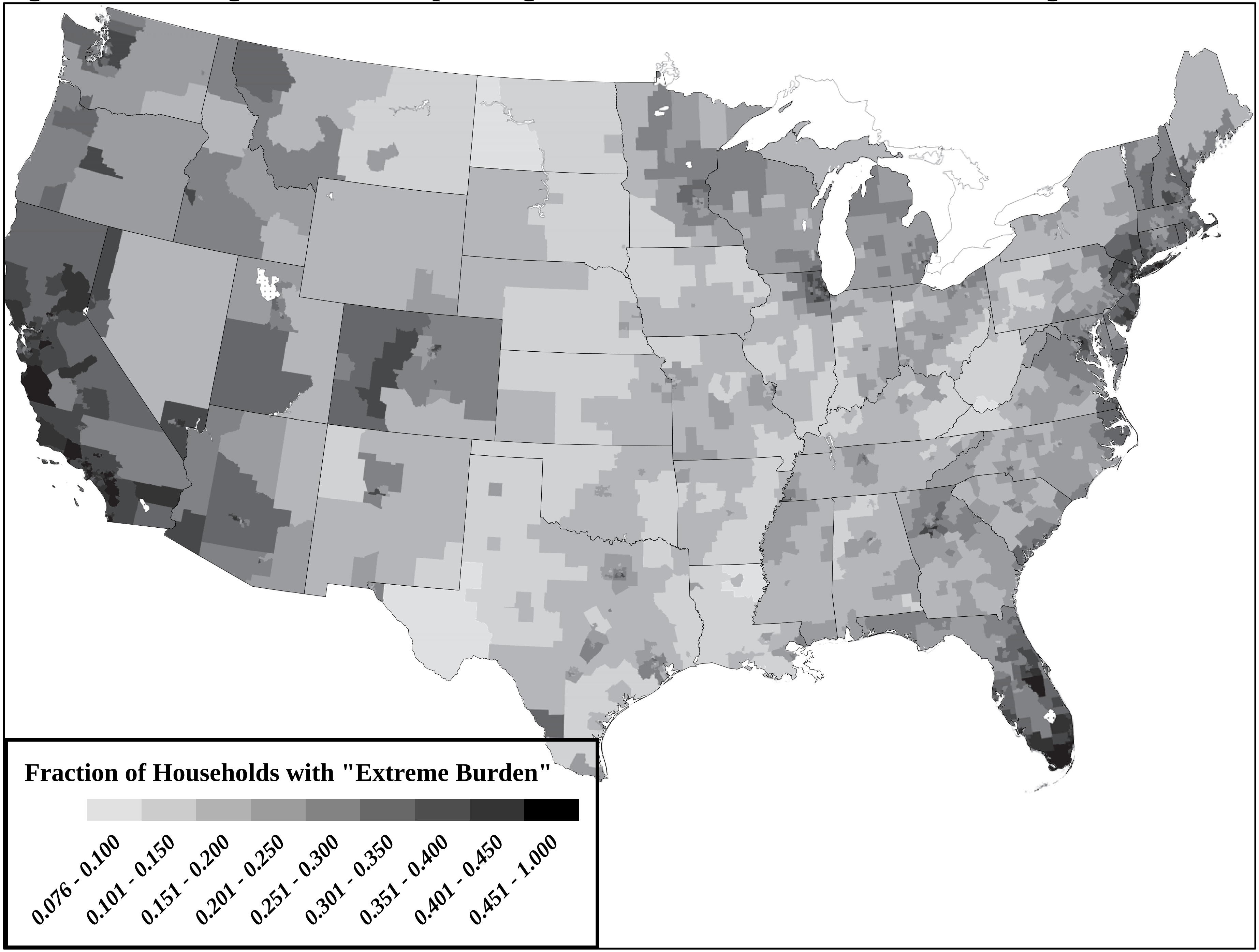


Figure 1C: Housing Unaffordability and Homeownership 1970-2013



Note: For non-BEA series, a moving average with weight of 0.5 for the years before and after shown. Housing shares include renters and owners; rent shares include renters only. BEA = Bureau of Economics Analysis, CEX = Consumer Expenditure Survey, AHS = American Housing Survey

Figure 2: Working Households Spending Over 50 Percent of Income on Housing, Percent: 2010



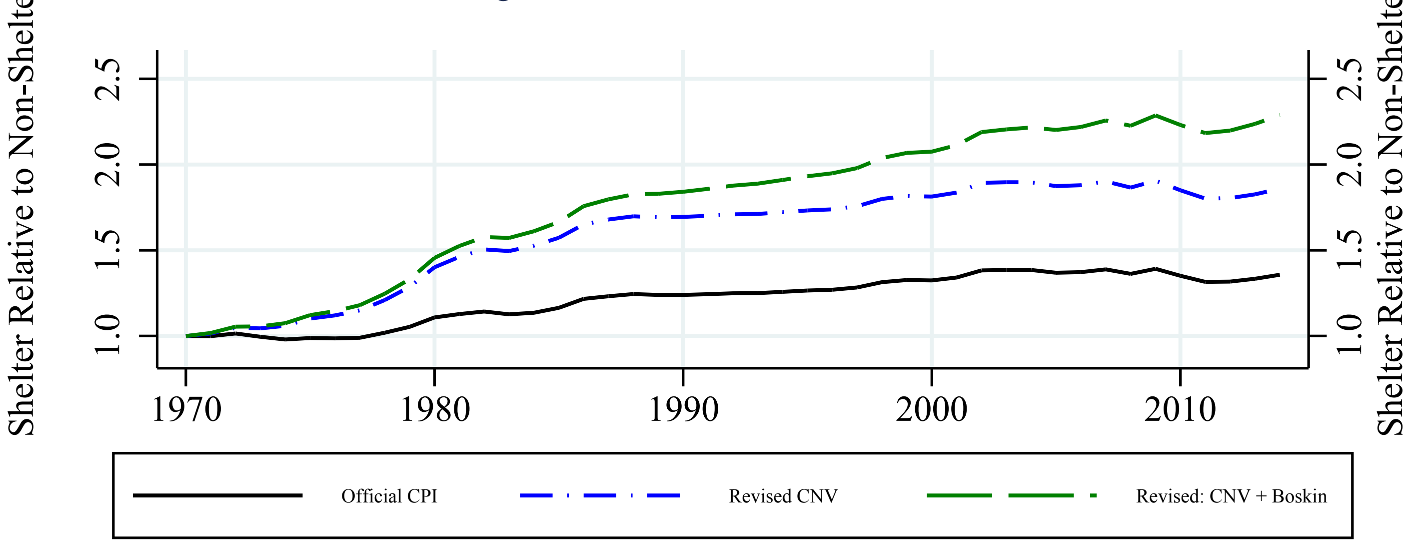


Figure 3A: The Relative Price of Shelter

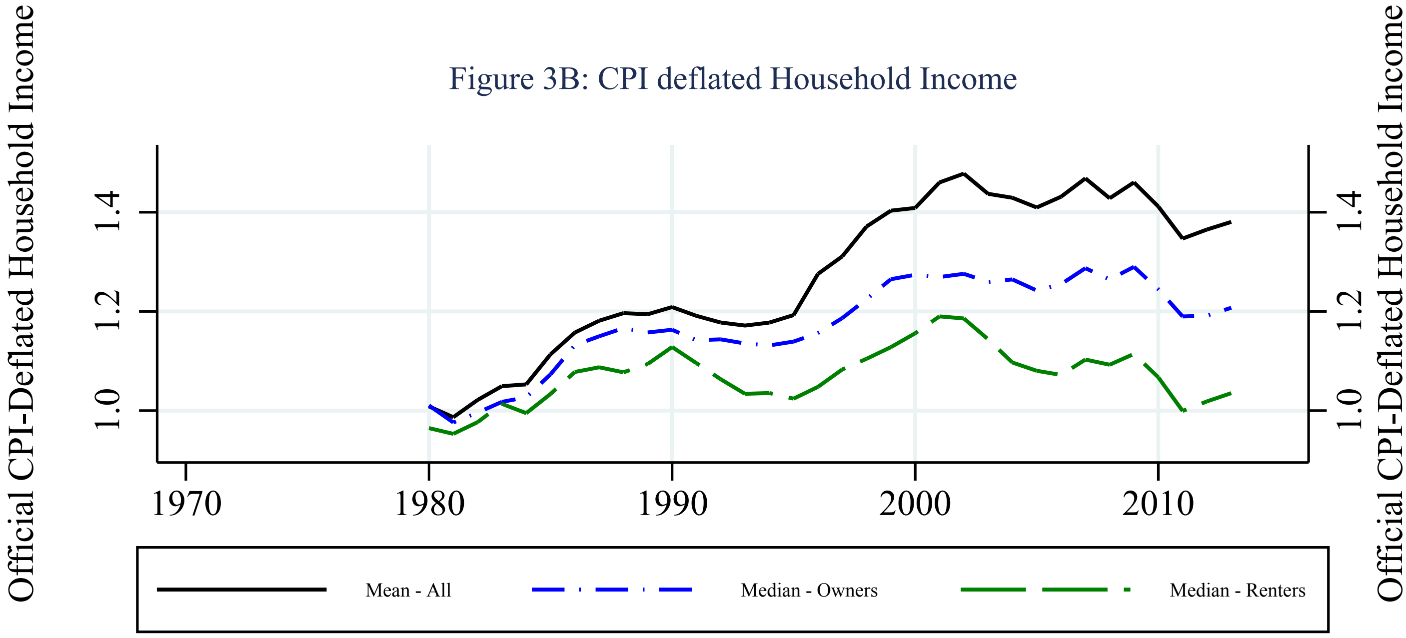


Figure 3B: CPI deflated Household Income

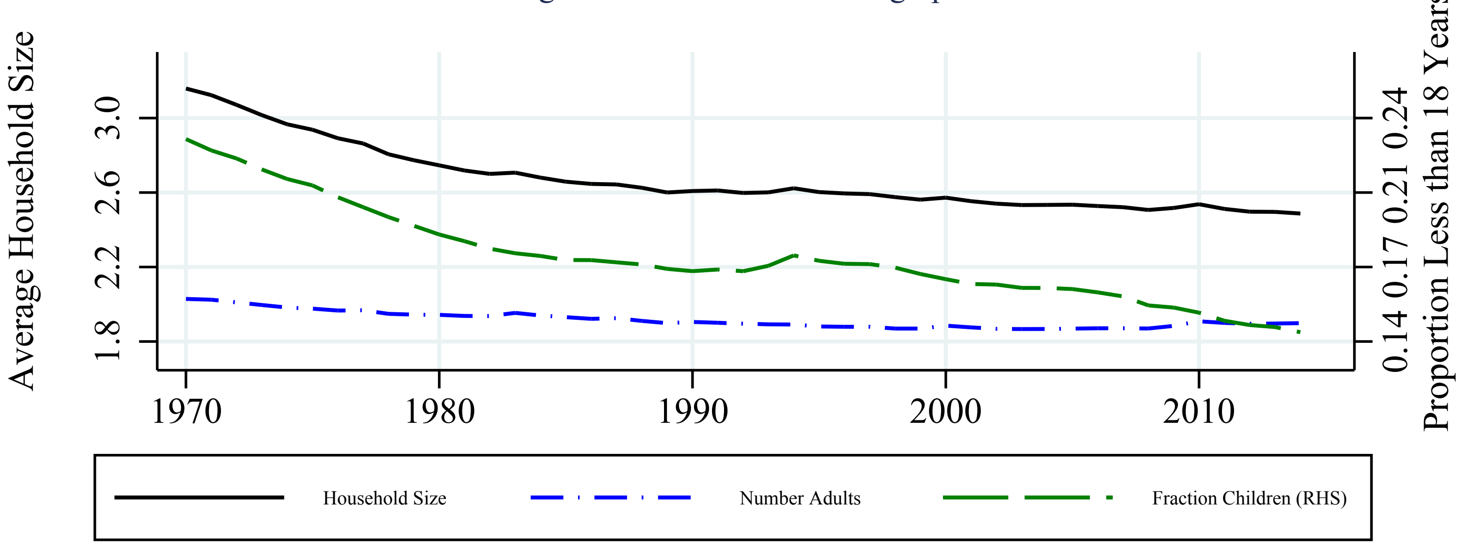


Figure 3C: Household Demographics

Note: CPI = Consumer Price Index. CNV = Crone, Nakamura, and Voith (2010), Boskin from New Product bias in Table 2 of the Boskin Commission (1996) Report. Income and Household Demographics from Current Population Survey. We normalize relative prices to 1 in 1970 and income to 1 in 1976 when renter status known.

Figure 4: Housing Consumption with Production Possibility Expansions

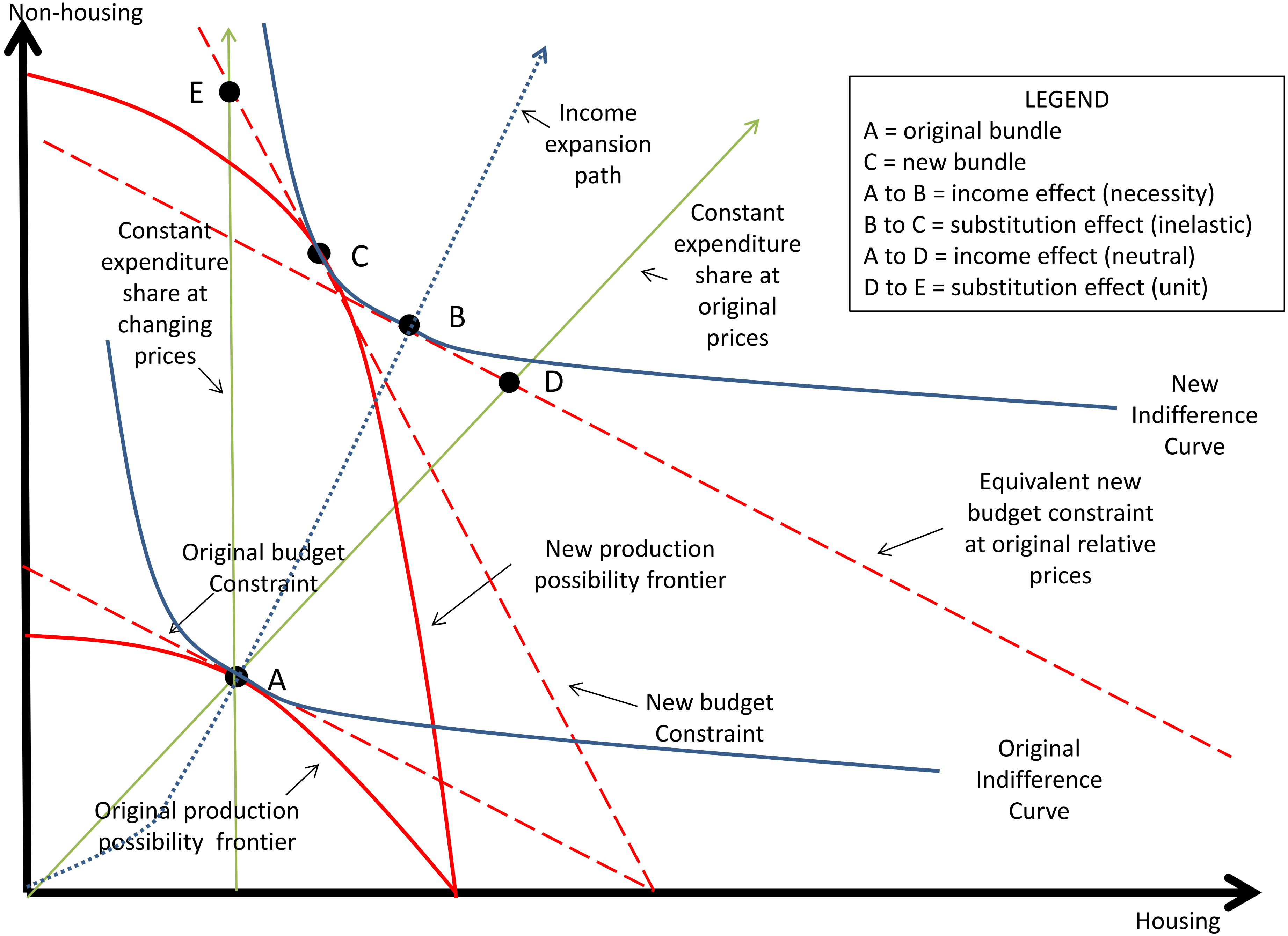


Figure 5A: Non-Housing Price Index vs. Housing Price Index, 2000

Figure 5B: Predicted Wage/Skill Index vs. Housing Price Index, 2000

Data Source: Non-housing price index is from CEO prices panel, housing price index is from Census

Figure 6A: Median Share of Income Spent on Rent
and the Relative Price of Housing, Renters Only 2000

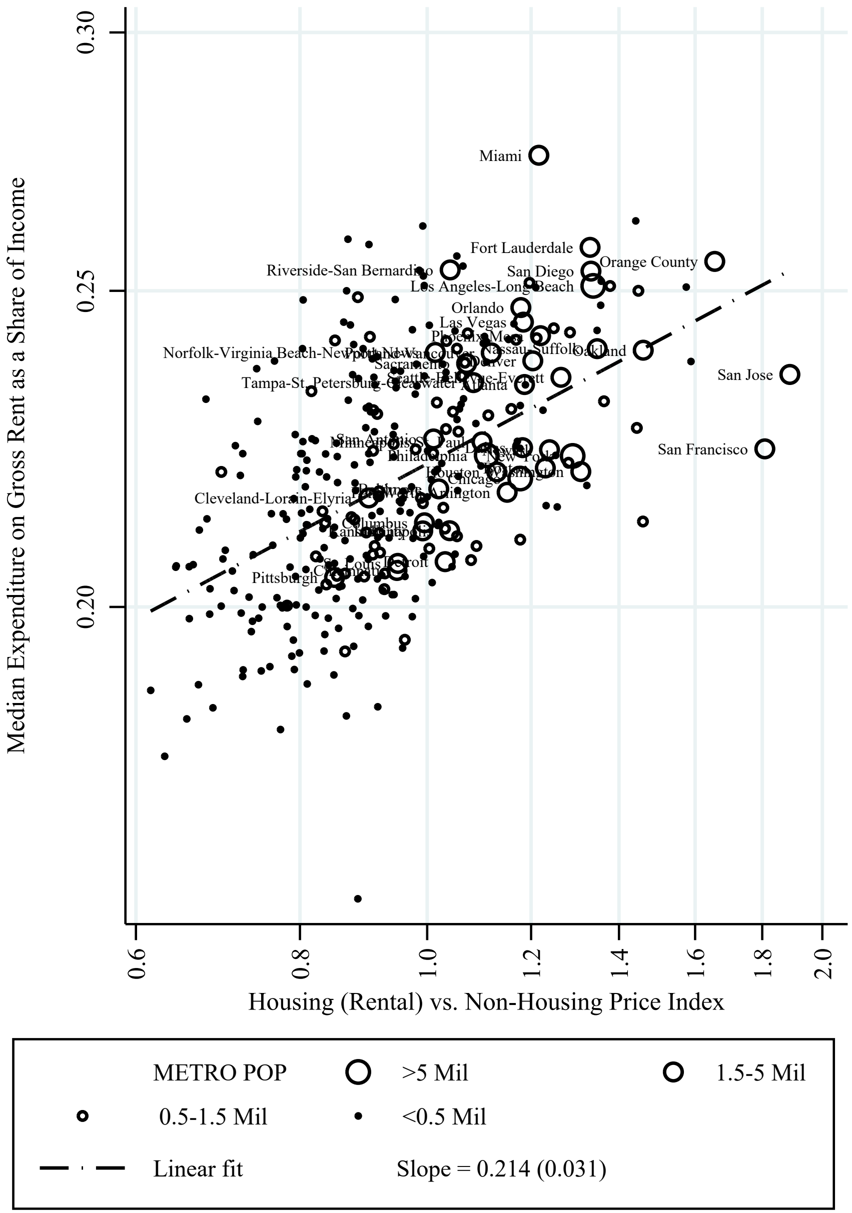


Figure 6B: Median Share of Income Spent on Housing
and the Relative Price of Housing, All Households 2000

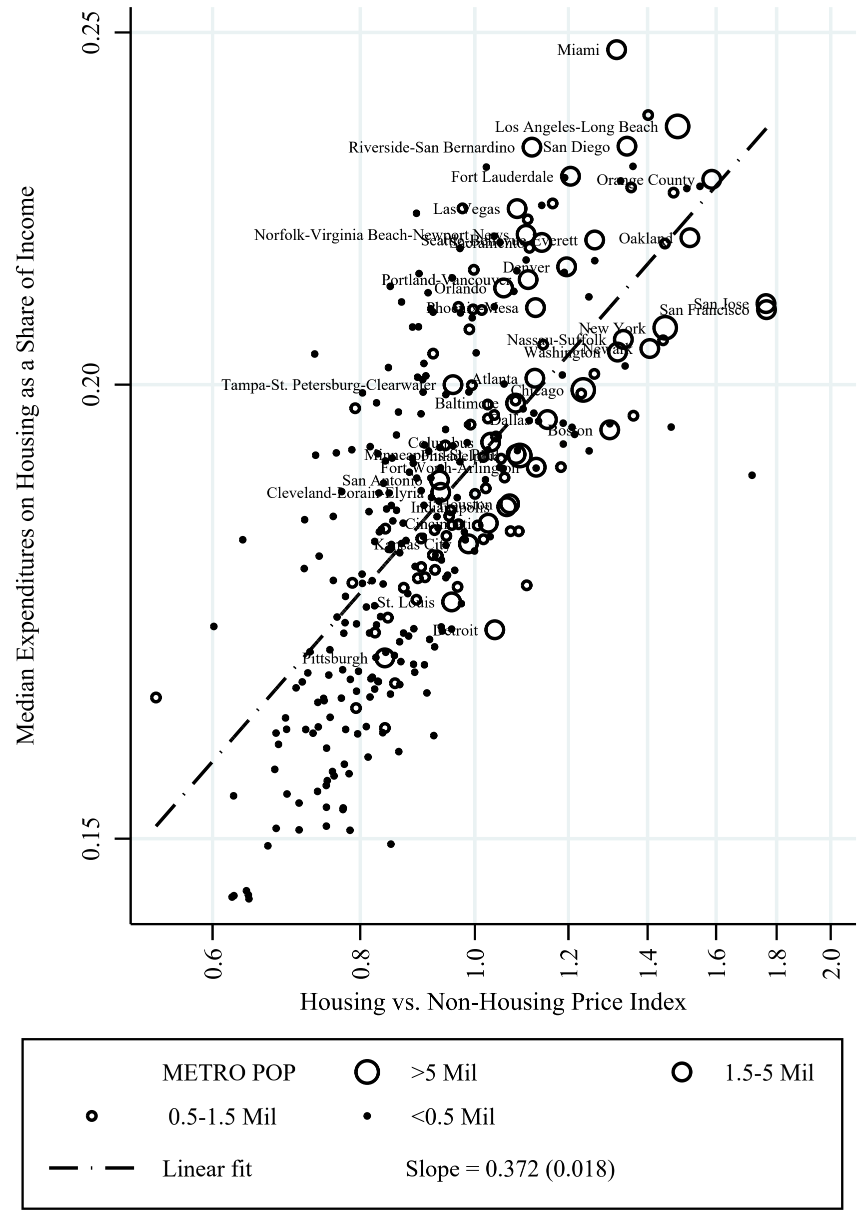


Figure 7: Comparison of Ideal Cost-of-Living Indices by Functional Form

Fig. 7A : Moderate substitution; housing necessity $\sigma = 2/3$, $\delta = 1/8$, $\varepsilon_{y,m} = 2/3$

Fig. 7B: Low substitution; housing weak necessity $\sigma = 1/2$, $\delta = 1/16$, $\varepsilon_{y,m} = 5/6$

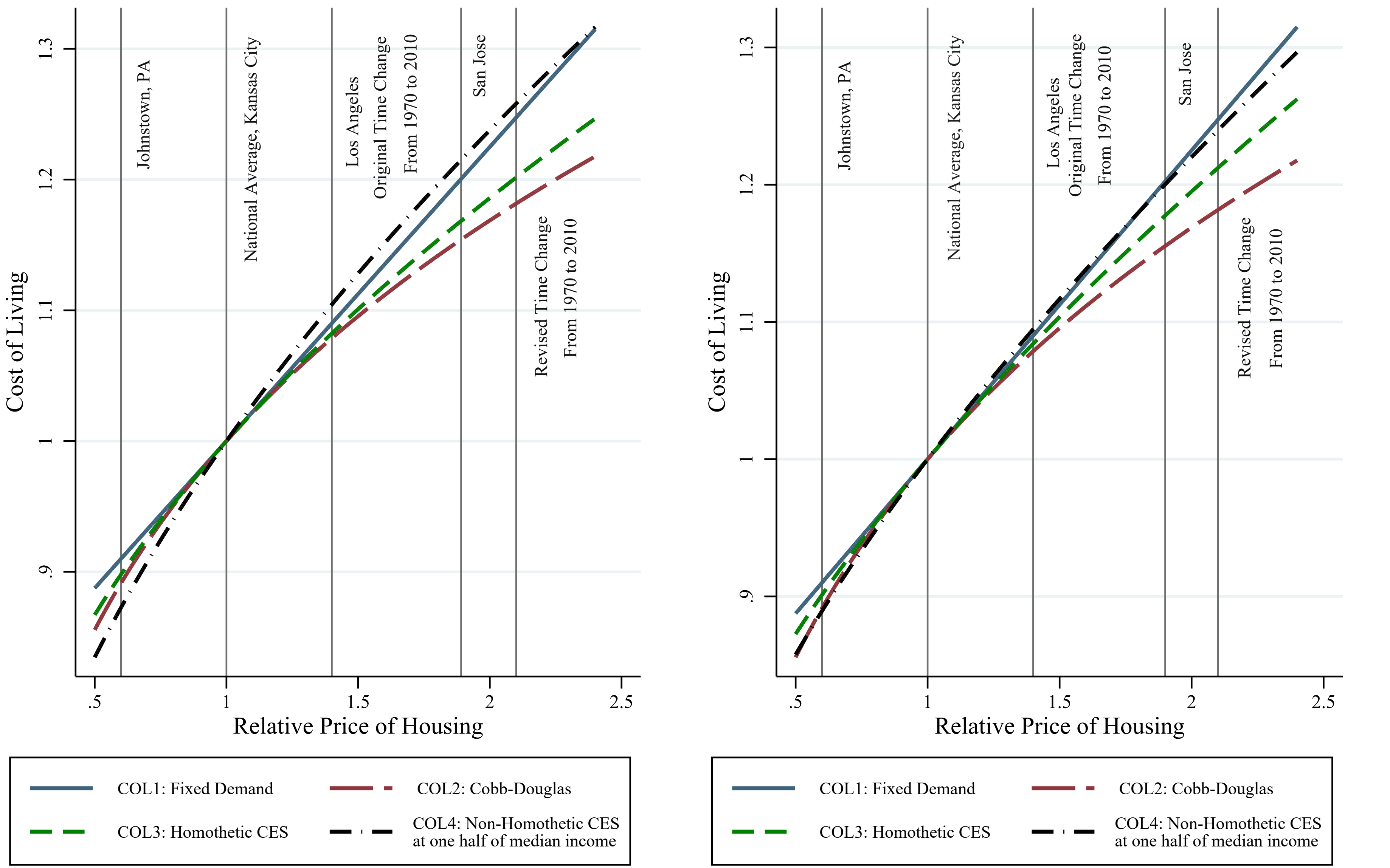
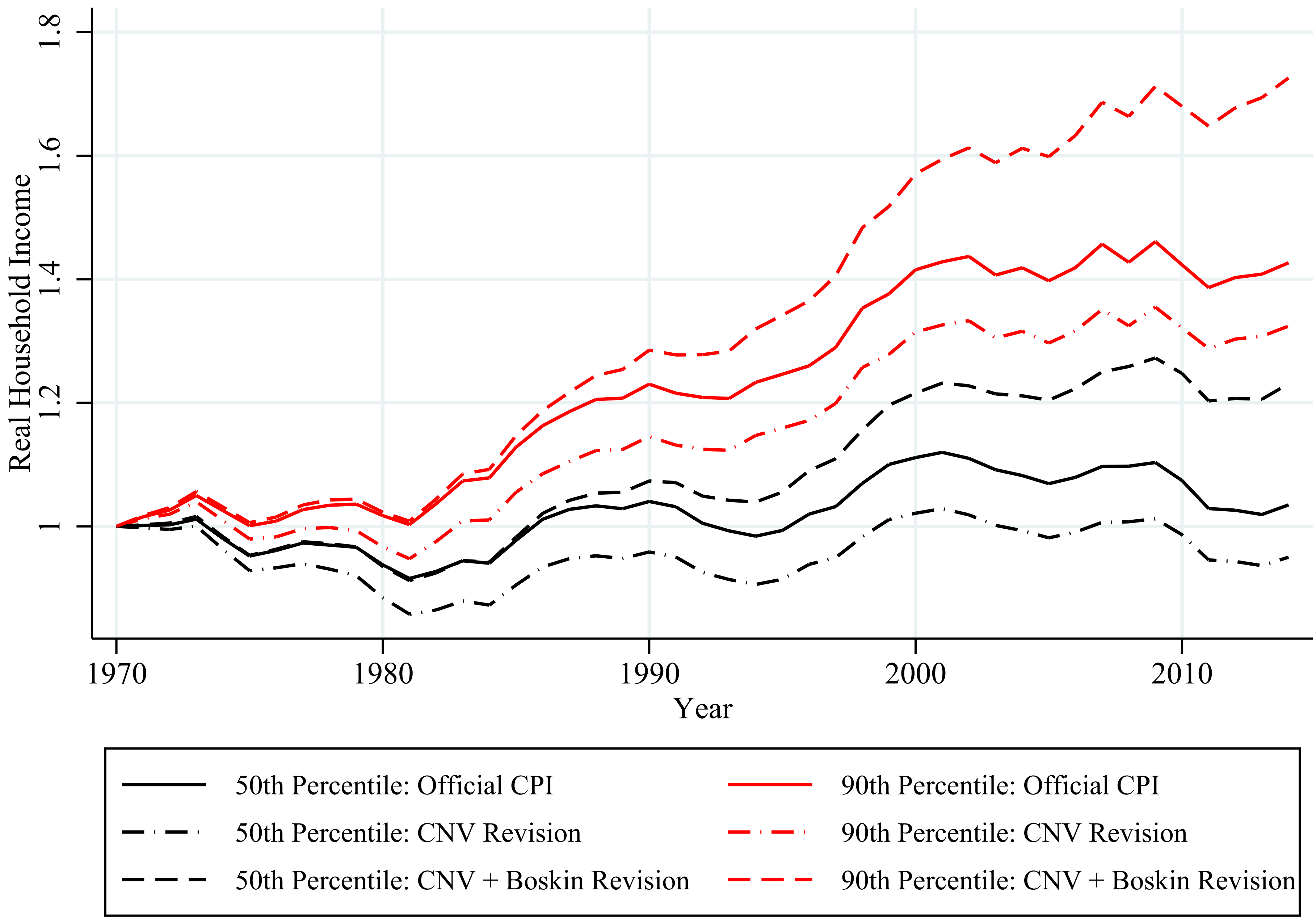


Figure 8: Real Household Income Deflated by Cost-of-living Index, 1970-2014



Note: Real household income at each percentile to 1 in 1970. Uses $\sigma = 2/3$, $\varepsilon_{y,m} = 2/3$

Appendix

A Separable Family of CES

A.1 Formulation and Parameters

We use the simple "separable family" of CES utility function from Sato (1977), and complement it with the Barten (1964) model as well as add a quality of life parameter:

$$U = Q\left[\frac{\delta_1(yn^{1-\phi})^\rho + \theta_1}{\delta_2 x^\rho + \theta_2}\right]^{\frac{1}{\rho\gamma}}$$

where $\theta_i = -\left(1/\gamma - \delta_i\right)\rho - \delta_i$ is composed of more elementary parameters. These are the distribution parameter, $\delta_1 = \delta$ and $\delta_2 = \delta - 1 < 0$, the substitution parameter, $\sigma = 1/(1-\rho)$, and the non-homotheticity parameter, γ. Raising the arguments by $1/\gamma$ helps with the limiting case as $\gamma \to 0$. The utility function is express in per–capita terms, so that co-habitating with others consuming the same amount contributes $n^{1-\phi}$ times the amount from sharing.

A.2 Marginal rate of substitution

Taking the ratio of partial derivatives, the marginal rate of substitution between the housing and non-housing goods is

$$MRS_{y,x} = \frac{\delta}{1-\delta}\left(\frac{x}{y}\right)^{1-\rho} n^{\rho(1-\phi)}\frac{\delta(yn^{1-\phi})^\rho + \theta_1}{(\delta-1)x^\rho + \theta_2} = \frac{\delta}{1-\delta}\left(\frac{y}{x}\right)^{\frac{1}{\sigma}}\left[\left(\frac{u}{Q}\right)^\gamma n^{1-\phi}\right]^{\frac{\sigma-1}{\sigma}}$$

At the household's optimal consumption bundle, $c/p = MRS_{x,y}$, implying:

$$\frac{x}{y} = \left(\frac{p}{c}\cdot\frac{1-\delta}{\delta}\right)^\sigma\left[\left(\frac{u}{Q}\right)^\gamma n^{1-\phi}\right]^{1-\sigma}$$

A.3 Expenditure Share on Housing

To solve for the expenditure share on housing, note that $d\ln y/d\ln x = dy/dx(x/y) = cx/py = s_x/s_y$. Then the ratio of the expenditure share spent on x to the share spent on y is:

$$\frac{s_x}{s_y} = \frac{\mathrm{d}\ln y}{\mathrm{d}\ln x} = \frac{cx}{py} = \left(\frac{1-\delta}{\delta}\right)^{\sigma}\left(\frac{c}{p}\right)^{1-\sigma}\left(\frac{u}{Q}\right)^{\gamma(1-\sigma)} n^{(1-\phi)(1-\sigma)}$$

With $\sigma < 1$ and $\gamma > 0$, the relative share of x to y increases with u/Q as well as c. It increases with n when $\phi < 1$. Then to solve for the housing expenditure share s_y, add one and invert:

$$\frac{1}{s_y} = \frac{cx}{py} + 1 \Rightarrow s_y = \frac{\delta^{\sigma}\left(\frac{p}{n^{1-\phi}}\right)^{1-\sigma}}{\delta^{\sigma}\left(\frac{p}{n^{1-\phi}}\right)^{1-\sigma} + (1-\delta)^{\sigma}c^{1-\sigma}\left(\frac{u}{Q}\right)^{\gamma(1-\sigma)}}.$$

Taking logarithms, we obtain an only partly linear equation

$$\ln s_y = \sigma\ln\delta + (1-\sigma)[\ln(p) - (1-\phi)\ln(n)] - \ln\left[\delta^{\sigma}\left(\frac{p}{n^{1-\phi}}\right)^{1-\sigma} + (1-\delta)^{\sigma}c^{1-\sigma}\left(\frac{u}{Q}\right)^{\gamma(1-\sigma)}\right]$$

To complete the log-linearization, we take the total derivative to get the approximation:

$$\widehat{s}_y = (1-s_y)(1-\sigma)\left[\hat{p} - \hat{c} - \gamma\hat{u} + \gamma\hat{Q} - (1-\phi)\hat{n}\right]$$

Relating the above equation to the regression equation ((5b)) we can set gives $\beta_0 = \sigma\ln\delta = \ln\bar{s}_y, \beta_1 = (1-\bar{s}_y)(1-\sigma)$, and $\gamma = -\beta_3/\beta_1$ if the term in brackets is set to one for reference prices $p = c = 1$ and a reference household size normalized to 1 $n = 1$ (so that n is household size relative to the mean). Then we need to set

$$\frac{u}{Q} = \left[\frac{1-\delta^{\sigma}}{(1-\delta)^{\sigma}}\right]^{\frac{1}{\gamma(1-\sigma)}}$$

For instance, suppose $e^{\beta_0} = 1/4$, and $\beta_1 a = 1/2$, then $\sigma = 2/3$ and $\delta = 1/8$. Then we need $u/Q = 0.64$ $\beta_3 = -\gamma(1 - \bar{s}_y)(1 - \sigma)$. The parameters can thus be expressed recursively as $\sigma = 1 - \beta_1/(1 - e^{\beta_0}), \delta = e^{\beta_0/\sigma}$

A.4 Hicksian Demand and Expenditure Functions

The Hicksian, or compensated, demands for the housing and non-housing goods associated with this utility function can be derived as:

$$y = \left(\delta \frac{\tilde{e}}{p} n^{1-\phi}\right)^{\sigma} \left[\theta_2 - \theta_1 \left(\frac{u}{Q}\right)^{\gamma \frac{1-\sigma}{\sigma}}\right]^{\frac{\sigma}{\sigma-1}} \tag{A.1}$$

$$x = \left[(1-\delta)\frac{\tilde{e}}{c}\right]^{\sigma} \left(\frac{u}{Q}\right)^{\gamma(1-\sigma)} \left[\theta_2 - \theta_1 \left(\frac{u}{Q}\right)^{\gamma \frac{1-\sigma}{\sigma}}\right]^{\frac{\sigma}{\sigma-1}} \tag{A.2}$$

where $\tilde{e}$ is a standard CES price index adjusted with an increasing utility weight on c and a division of p, so that

$$\tilde{e} = \tilde{e}\left[\frac{p}{n^{1-\phi}}, c\left(\frac{u}{Q}\right)^{\gamma}\right] = \left[\delta^{\sigma}\left(\frac{p}{n^{1-\phi}}\right)^{1-\sigma} + (1-\delta)^{\sigma} c^{1-\sigma}\left(\frac{u}{Q}\right)^{\gamma(1-\sigma)}\right]^{\frac{1}{1-\sigma}} \tag{A.3}$$

The expenditure for the NH-CES function with the Barten adjustment is then:

$$e(p, c, u, n; Q) = \left[\delta^{\sigma}\left(\frac{p}{n^{1-\phi}}\right)^{1-\sigma} + (1-\delta)^{\sigma} c^{1-\sigma}\left(\frac{u}{Q}\right)^{\gamma(1-\sigma)}\right]^{\frac{1}{1-\sigma}} \left[\theta_2 - \theta_1\left(\frac{u}{Q}\right)^{\gamma\frac{1-\sigma}{\sigma}}\right]^{\frac{\sigma}{\sigma-1}}. \tag{A.4}$$

B Data

We define cities at the Metropolitan Statistical Area (MSA) level using the 1999 Office of Management and Budget definitions of primary MSAs (e.g., San Francisco is separate from Oakland

and San Jose), of which there are 311. Data from the U.S. Census data from Integrated Public-Use Microdata Series (IPUMS), from Ruggles et al. (2004), for several purposes.

B.1 Wage Differentials

The wage differentials are calculated for workers ages 25 to 55 who report working at least 30 hours a week, 26 weeks a year. The MSA assigned to a worker is determined by their place of residence, rather than their place of work. The wage differential of an MSA is calculated by regressing log hourly wages on a rich set of covariates and a set of indicators for which MSA a worker lives in. The wage differentials are taken to be the coefficients on these MSA indicators, renormalized to have a national average value of zero. The covariates consist of:

- 12 indicators of educational attainment;
- a quartic in potential experience, and potential experience interacted with years of education;
- 9 indicators of industry at the one-digit level (1950 classification);
- 9 indicators of employment at the one-digit level (1950 classification);
- 4 indicators of marital status (married, divorced, widowed, separated);
- an indicator for veteran status, and veteran status interacted with age;
- 5 indicators of minority status (Black, Hispanic, Asian, Native American, and other);
- an indicator of immigrant status, years since immigration, and immigrant status interacted with black, Hispanic, Asian, and other;
- 2 indicators for English proficiency (none or poor).

All covariates are interacted with gender.

This regression is run using census-person weights.

B.2 Housing Rent and Price Indices

The housing rent and price differentials are calculated using the logarithm of rents, whether they are reported gross rents or imputed rents derived from housing values. The differential housing price of an MSA is calculated in a manner similar to the wage differential, except using a regression of the actual or imputed rent on a set of covariates at the unit level and a set of MSA indicators. The covariates for the adjusted differentials are:

- 9 indicators of building size;
- 9 indicators for the number of rooms, 5 indicators for the number of bedrooms, number of rooms interacted with number of bedrooms;
- 2 indicators for lot size;
- 7 indicators for when the building was built;
- 2 indicators for complete plumbing and kitchen facilities;
- 8 indicators for home heating fuel;
- an indicator for commercial use;
- an indicator for condominium status (owned units only).

We first run a regression of housing values on housing characteristics and MSA indicator variables weighting by census-housing weights. The housing-price index are taken from the MSA indicator variables in this regression, renormalized to have a national average of zero.

B.2.1 Alternative Census Housing Price Index

The Alternative Census Housing Price Index are estimated from the 2000 united States Census 5% data from the Integrated Public-Use Microdata Series (IPUMS), following Malpezzi, Chun and

Green (1998). The housing price differentials are calculated using the logarithm of rents, whether they are reported gross rents or imputed rents derived from housing values. We first fit separate regressions for each MSA, regressing the log yearly rents on a set of MSA dummies and a number of covariates at the unit level. We then use the predicted price from each regression in each location to get the normalized price index. The covariates for the adjusted differential are:

- 9 indicators of building size;
- 9 indicators for the number of rooms, 5 indicators for the number of bedrooms, number of rooms interacted with number of bedrooms;
- 2 indicators for lot size;
- 7 indicators for when the building was built;
- 2 indicators for complete plumbing and kitchen facilities;
- 8 indicators for home heating fuel;
- an indicator for commercial use;
- an indicator for condominium status (owned units only).

We first run a hedonic regression for each MSA, using housing characteristics alone. Second, we calculate predicted housing prices in each MSA from each regression, and calculate the MSA-level means. Third, we obtain the normalized housing price index for each MSA by using the predicted values of housing minus the national average.

B.2.2 CEX Housing Price Index

The CEX Housing Price Index is computed from 1997-2003 pooled Consumer Expenditure Survey. The housing price differentials are calculated using the logarithm of rents, whether they are reported gross rents or imputed rents derived from housing values. We regress the log yearly rents on a set of geographical area dummies and a number of covariates at the unit level. The geographical area is defined based on state, population size, and whether it is in a metro area. In order to compare with the other price indices, we match CEX geographical units with Census PMSAs by state, population, and metropolitan area status. The matching process is not perfect, since a state may have two MSAs with indistinguishable populations, preventing us from differentiating them.

The covariates for the adjusted differentials are:

- 9 indicators of building size;
- 9 indicators of building structure;
- 9 indicators for the number of rooms, 5 indicators for the number of bedrooms, number of rooms interacted with number of bedrooms;
- 5 indicators for the number of bathrooms;
- 7 indicators for when the building was built;
- 2 indicators for complete plumbing and kitchen facilities;
- 4 indicators for home heating fuel;
- an indicator for commercial use;
- an indicator for condominium status (owned units only).

We first run a regression of housing values on housing characteristics and geographical area indicator variables weighting by CEX-housing weights. The housing-price index is taken from the coefficients on the geographical area indicator variables in this regression, renormalized to have a national average of zero.

B.2.3 CEO Prices Panel Housing Price Index

We use the Carrillo, Early, Olsen (2013) Prices Index Panel for all areas in the United States in the year 2000. CEO's source of housing data is HUD's 2000 Section 8 Customer Satisfaction Survey (CSS). They produce a geographic housing price index for 2000 by estimating a hedonic regression. They regress the logarithm of gross rents on observed characteristics of the rental units and their neighborhoods, other determinants that reflect unobserved characteristics that affect market rents, and a set of geographic area dummies for metropolitan areas and the non-metropolitan areas of each state.

B.3 Housing Expenditure Share

B.3.1 Census Housing Expenditure Share

The Census housing expenditure share is calculated from the 2000 United States 5% data from the Integrated Public-Use Microdata Series (IPUMS). The housing expenditure share is calculated as the ratio of housing expenditure to household income. For renters, we use gross rent as housing expenditure, while for owners, we use imputed rents derived from housing values plus utility fees. The cross-MSA mean of the MSA-level median rental share is .225 and the mean of the MSA-level meadian housing share for both renters and owners is 0.196.

B.3.2 AHS Rental Share

The AHS rental share is computed from the 1974-2013 American Housing Survey microdata. The AHS housing expenditure share is defined as the ratio of monthly housing cost to household income.

B.3.3 CEX Rental Share

The CEX rental share in appendix table 2 is derived from 1997-2003 Consumer Expenditure Survey microdata. The rental share in figure 1B is derived from 1974-2014 microdata. The CEX rental share is computed as the ratio of expenditure on rents to total expenditure. We define geographies in the CEX as discussed in section B.2.2.

APPENDIX TABLE 1: COMPENSATED DEMAND FUNCTIONS - ADDITIONAL YEARS, DATASETS, AND PRICE INDICES - RENTERS ONLY

Dataset/Price Index:	Census 1980	Census 1990	ACS 2007-11	Alt Housing Price Index	CEO Housing Price Index
	(1)	(2)	(3)	(4)	(5)
Dependent Variable:	Log Median Rental Share				
Panel A: Regression Results					
Rental/Housing Price Index	0.219	0.246	0.283	0.203	0.118
	(0.042)	(0.028)	(0.028)	(0.027)	(0.034)
Non Housing Price Index	-0.219	-0.246	-0.283	-0.203	-0.118
	(0.042)	(0.028)	(0.028)	(0.027)	(0.034)
Predicted Wage Index	-0.407	-0.266	-0.483	-0.235	-0.163
	(0.133)	(0.141)	(0.102)	(0.115)	(0.134)
Constant	-1.538	-1.476	-1.275	-1.492	-1.492
	(0.006)	(0.007)	(0.005)	(0.005)	(0.005)
Sample Size (number of areas)	328	331	331	331	331
Adjusted R-squared	0.166	0.604	0.582	0.481	0.379
Homogeneity of Demand Restricted	Yes	Yes	Yes	Yes	Yes
Unconstrained Sum of Housing and Non-Housing Price Index Coefficients	-0.175	-0.355	0.149	0.058	0.146
	(0.207)	(0.173)	(0.068)	(0.073)	(0.098)
P-value of Test of Homogeneity of Demand	0.398	0.041	0.030	0.429	0.137
Panel B: Implied Demand Parameters					
Geometric Mean Expenditure Share	0.215	0.229	0.279	0.225	0.225
	(0.001)	(0.002)	(0.001)	(0.001)	(0.001)
Uncompensated Own Price Elasticity of Housing Demand	-0.693	-0.694	-0.582	-0.744	-0.846
	(0.053)	(0.051)	(0.047)	(0.043)	(0.057)
Income Elasticity of Housing Demand	0.593	0.734	0.517	0.765	0.837
	(0.133)	(0.141)	(0.102)	(0.115)	(0.134)
Elasticity of Substitution Between Housing and Consumption Goods σ	0.720	0.682	0.607	0.738	0.848
	(0.054)	(0.037)	(0.039)	(0.035)	(0.044)
Distribution Parameter δ	0.882	0.885	0.878	0.867	0.828
	(0.018)	(0.013)	(0.016)	(0.012)	(0.015)
Non-homotheticity Parameter γ	1.856	1.083	1.705	1.157	1.382
	(0.682)	(0.537)	(0.333)	(0.534)	(0.980)

All specifications include renters only. Columns 1 through 3 use the 1980 Census, 1990 Census, and 2007-2011 American Community Surveys to calculate house price indices and rental shares as in table 1. Column 4 calculates the house price index similarly to Malpezzi et al. (1998) as described in appendix B.2.1. Column 5 constructs the house price index using the Price Indexes Panel provided by Carrillo, Early, and Olsen (2013), described in appendix B.2.4. Columns 4 through 5 use the share of income devoted to housing from the 2000 Census as the rental share.

APPENDIX TABLE 2: METRO LEVEL REGRESSIONS - CONSUMER EXPENDITURE SURVEY

Household Type	Renters only	Renters and Owners	Renters only	Renters and Owners
	(1)	(2)	(3)	(4)
Regression Type:	Compensated	Compensated	Uncompensated	Uncompensated
Panel A: Regression Results				
Housing Price Index	0.341	0.525	0.685	0.723
	(0.056)	(0.042)	(0.075)	(0.077)
Log Median Predicted Expenditure	-0.086	0.001	-0.644	-0.549
	(0.164)	(0.129)	(0.075)	(0.081)
Constant	-1.013	-1.174	-1.013	-1.174
	(0.008)	(0.007)	(0.006)	(0.006)
Sample Size (number of areas)	163	163	163	163
Adjusted R-squared	0.289	0.573	0.723	0.815
P-value of Test of Homogeneity of Demand	0.000	0.000	0.736	0.004
Panel B: Implied Demand Parameters				
Geometric Mean Expenditure Share	0.363	0.309	0.363	0.309
	(0.003)	(0.002)	(0.002)	(0.002)
Uncompensated Own Price Elasticity of Housing Demand	-0.628	-0.476	-0.315	-0.277
	(0.104)	(0.072)	(0.075)	(0.077)
Income Elasticity of Housing Demand	0.914	1.001	0.356	0.451
	(0.164)	(0.129)	(0.075)	(0.081)
Elasticity of Substitution Between Housing and Consumption Goods σ	0.465	0.241	0.292	0.199
	(0.089)	(0.061)	(0.083)	(0.081)

The data are from Consumer Expenditure Survey 1997-2003.All regressions are constrained to exhibit homogeneity of demand. Standard errors are clustered at the metro level. The predicted income measure in Panel A is the median predicted expenditures of all adults in the household.

Figure A: Alternative vs. Census Housing Price Index

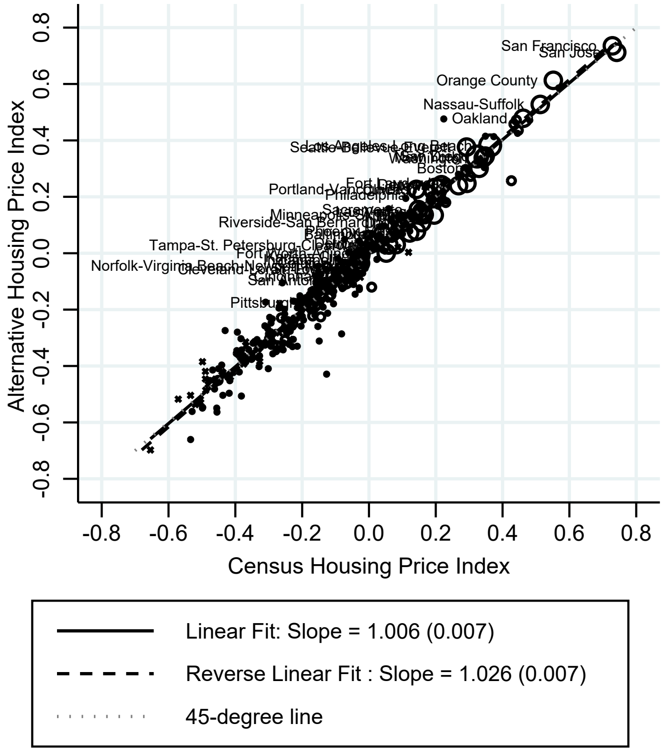


Figure B: CEO vs. Census Housing Price Index

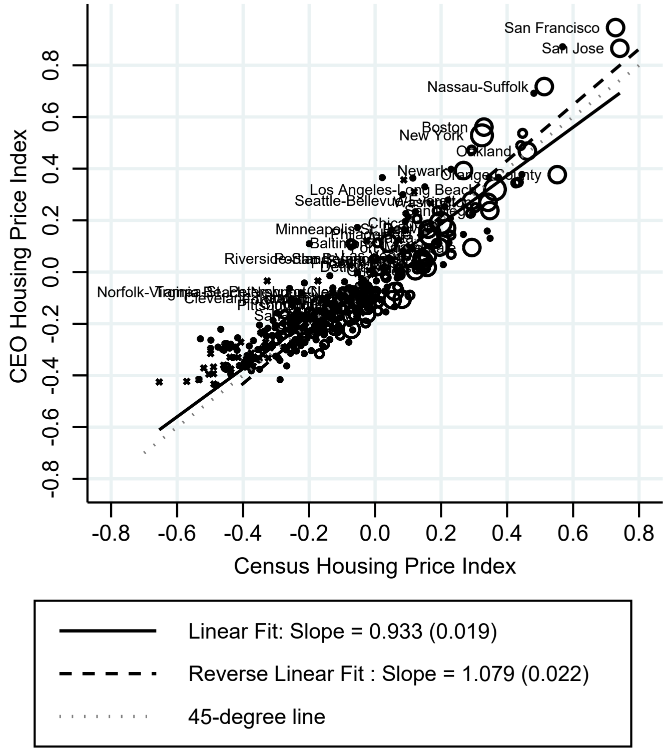


Figure C: AHS vs. Census Housing Price Index

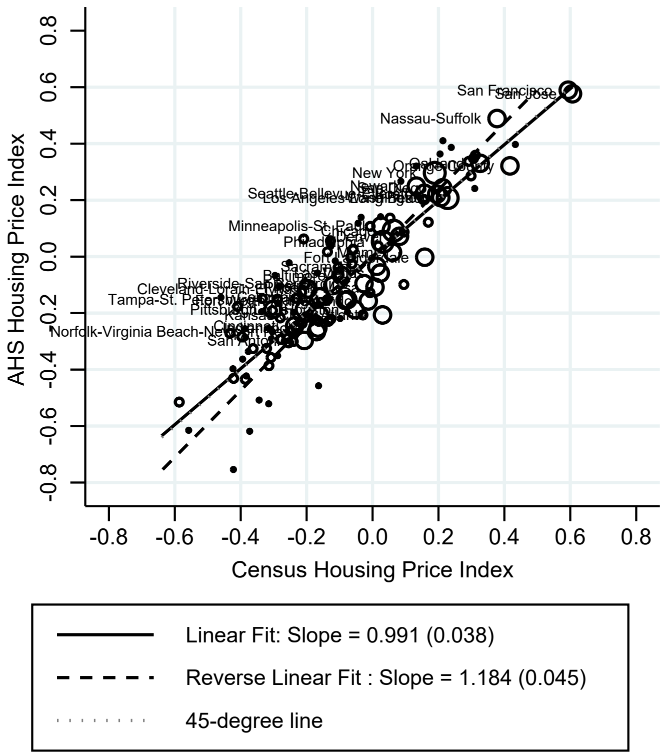
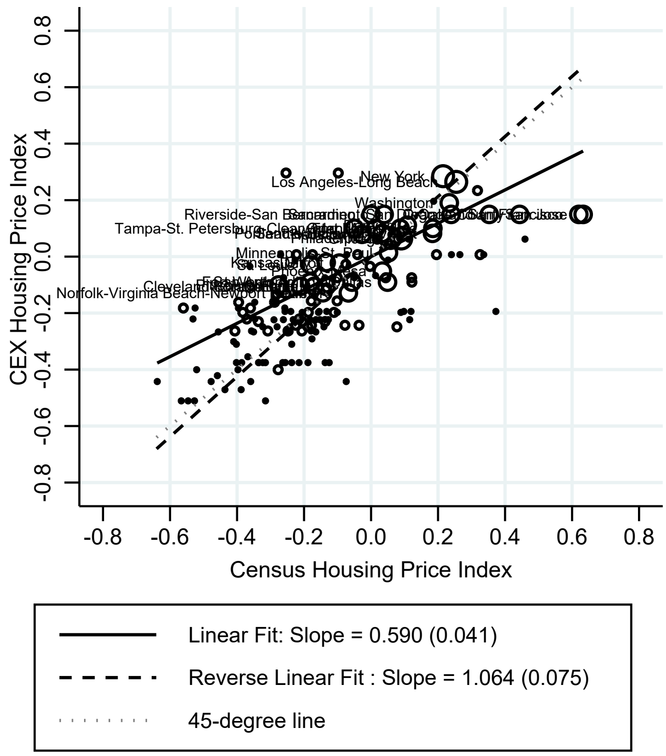


Figure D: CEX vs. Census Housing Price Index



Data Source: 2000 Census, 1997-2003 CEX, 1997-2003 AHS, renters only