

# Income Distribution and Housing Prices: An Assignment Model Approach

Niku Määttänen  
ETLA and HECER\*

Marko Terviö  
Aalto University and HECER<sup>†</sup>

December 9, 2013<sup>‡</sup>

## Abstract

We present a framework for studying the relation between the distributions of income and house prices that is based on an assignment model where households are heterogeneous by incomes and houses by quality. Each household owns one house and wishes to live in one house; thus everyone is potentially both a buyer and a seller. In equilibrium, the distribution of prices depends on both distributions in a tractable but nontrivial manner. We show how the impact of increased income inequality on house prices depends on the shapes of the distributions, and can be inferred from data. In our empirical application we find that increased income inequality between 1998 and 2007 had a negative impact on average house prices in 6 US metropolitan areas. JEL: D31, R21.

---

\*niku.maattanen@etla.fi, The Research Institute of the Finnish Economy (ETLA) and Helsinki Center for Economic Research.

<sup>†</sup>marko.tervio@aalto.fi, Aalto University School of Economics and Helsinki Center for Economic Research.

<sup>‡</sup>We thank Essi Eerola, Pauli Murto, Ofer Setty, Otto Toivanen, Juuso Välimäki, and seminar participants at Bonn, ECARES, EIEF, EUI, Mannheim, Munich, NHH Bergen, NorMac, Nottingham, Stockholm, SOFI, UAB, and Zurich for useful suggestions. Määttänen thanks Suomen arvopaperimarkkinoiden edistämissäätiö and the Academy of Finland and Terviö thanks the European Research Council for financial support.

# 1 Introduction

A central feature of the housing market is that housing is not a fungible commodity but comes embedded in indivisible and heterogeneous units. What people refer to as “houses,” are really bundles of land and structures. The quality of land is inherently heterogeneous because locations differ in their attractiveness due to fixed factors such as distance from the center and view of the sea. The supply of structures is more or less fixed in the short term, and only partly adjustable in the medium term. Quality of structures can also have a fixed component, due to zoning restrictions or the scarcity value of vintage architecture. At the same time, housing is a normal good, as the more wealthy spend more on housing.

Given the indivisibility and heterogeneity of houses, it is natural to consider the housing market as an assignment problem. Yet, assignment models are not a standard tool in applied work on housing economics. One reason may be that in the standard assignment model agents are *ex ante* divided into two distinct classes, such as "buyers" and "sellers". The two-sidedness assumption is natural in many markets, for example when firms are matched with workers, but is problematic in the housing market where most households are potentially both sellers and buyers. Furthermore, standard assignment models assume transferable utility and so preclude income effects. In the housing market, however, income effects may be significant since housing takes up a large part of household expenditure.

In this paper, we present a one-sided assignment model with non-transferable utility, where houses are heterogeneous by quality and households are heterogeneous by income. Each household owns one house and wishes to live in one house; thus households' reservation prices as sellers depend on the opportunities available to them as buyers. We model a single metropolitan region, where the set of households is fixed. The distributions of income and house quality are exogenous, while the distribution of house prices and therefore wealth is endogenous, with the exception of the cheapest or “marginal” house. In general, the initial joint distribution of houses and income is arbitrary, which results potentially in a lot of trading between households. Equilibrium prices depend on the properties of the joint distribution, not just on the marginal distributions of income and house quality.

We show the existence of equilibrium prices under an arbitrary initial endowment. For most of the paper we focus on the equilibrium prices that emerge after all trading opportunities have been exploited. Equilibrium prices can then be solved explicitly from the joint distribution of wealth and house quality and the parameters of the utility function. We use the equilibrium conditions to derive analytical results for the impact of income distribution on the distribution of house prices.

According to our theoretical results the impact of increased income inequality on top (as well as average) house prices is ambiguous. The intuition for why increase in income inequality leads to lower prices at the bottom of the quality distribution is clear: if low-income households have less income they bid less for low-quality houses. However, in equilibrium, any changes in prices spill upwards in the quality distribution. This is because the binding outside opportunity of any (inframarginal) household is that they must want to buy their equilibrium match rather than the next best house. The equilibrium price gradient—the price difference between two "neighboring" houses in the quality distribution—is pinned down by how much the households at the relevant part of the income distribution are willing to pay for the quality difference. The price of any particular house is then given by the summation of all price gradients below, plus the price of the marginal house. Thus, while an increase in incomes at the top increases the local price gradient, the lower prices at the bottom put downward pressure on prices above. It is therefore possible for all house prices to go down in response to an increase in inequality.

In order to illustrate the mechanisms of the model and to evaluate their quantitative importance we apply our model to data from the American Housing Survey (AHS) in 1998 and 2007. We assume that all households prefer to live in their current house under current market prices. Under this "post-trade" interpretation of data the distribution of unobserved house qualities can be inferred as the distribution that gives rise to the observed price distribution as the equilibrium outcome of our model. The inferred distribution of house qualities can be used to generate counterfactuals. This setup is at its most useful in analyzing changes or policies that may be at least partly capitalized into house prices, such as income transfers and housing cost subsidies; we discuss other potential applications of our model in the end.

We find that a suitably parametrized constant elasticity of substitution (CES) utility function allows us to roughly match the observed change in the price distribution under the assumption that it was caused by the observed change in incomes while unobserved house qualities remained unchanged. We first use the calibrated model to illustrate how the income elasticity of housing demand depends on the quality and income distributions, and varies between households, despite homogeneous CES preferences. Intuitively, how much more one would spend on housing following an increase in income depends on the additional housing quality that one extra dollar can buy; this in turn varies over the distribution, as it depends both on available qualities and on the incomes of competing buyers and sellers.

We then consider counterfactual income distributions for 2007 where all incomes grow uniformly since 1998 at the same rate as the actual mean income in the same metropolitan

region. We compare house prices that result from the counterfactual income distributions with the prices resulting from the actual income distributions, while holding constant the population and the housing supply in each city. This exercise can be seen as a way to disentangle the short or medium run impact of increased income inequality on house prices from changes in other house price determinants.

Depending on the region, counterfactual income distributions result in house prices that are on average 0 – 10% higher than house prices resulting from actual income distributions. (This excludes any changes in the top 3% of the price distribution, which is excluded due to top coding). This implies that the increase in inequality has resulted in lower prices on average than would have prevailed under uniform income growth, but the impact is modest in magnitude compared to the overall change in prices. The contribution of uneven income growth on house prices has been positive only within the top decile, with magnitudes of up to 12%. We also compare our results to those that one would obtain using a two-sided model or a much simpler model without indivisibilities.

In the next section we discuss related literature. In Section 3 we present the model and our theoretical results. In Section 4 we show how the model can be used for inference and counterfactuals. The empirical application is presented in Section 5, and Section 6 concludes.

## 2 Related Literature

Assignment models are models of perfectly competitive matching markets that focus on the combined impact of indivisibilities and two-sided heterogeneity; for a review see Sattinger (1993). All other frictions, such as imperfect information or transaction costs, are assumed away. Both sides of the market are assumed to have a continuum of types, so there is no market power or bargaining as all agents have arbitrarily close competitors. Assignment models typically include an assumption of a complementarity in production, which results in assortative matching. Resulting equilibrium prices depend on the shapes of the type distributions on both sides of the market but in a reasonably tractable way. Assignment models have usually been applied to labor markets, where the productive complementarity is between job types and worker types, as in Sattinger (1979) and Teulings (1995).

There exists a large literature on two-sided assignment, where agents are ex ante divided into two distinct classes, such as "buyers" and "sellers," are matched.<sup>1</sup> A two-

---

<sup>1</sup>For recent theoretical literature see, for example, Legros and Newman (2007) and Caplin and Leahy (2010).

sided variant of our model can be obtained by assuming that houses are owned by a separate class of agents who do not themselves consume houses; it is essentially the setup analyzed by Braid (1981), to which we return in section 3.4. Our model differs from previous assignment models by featuring both one-sided matching and non-transferable utility. There is no complementarity in the usual sense, but equilibrium nevertheless involves assortative matching by wealth and house quality, essentially because housing is a normal good. We don't restrict the shapes of the distributions (beyond minimal regularity conditions). In the empirical application we can then use a nonparametric method to infer the unobserved type distribution and to construct counterfactuals similar to Terviö (2008).

Assignment models are rarely used in applied work on housing economics. An interesting exception is Landvoigt, Piazzesi, and Schneider (2012). They use a two-sided assignment model with concave utility to motivate a nonlinear relationship between housing quality and house prices, which they incorporate into an intertemporal model of housing demand. They employ a calibrated version of their model to study the housing market of San Diego county, California, during the 2000s boom. Using the model, they infer the role of cheap credit from the cross section of capital gains by house quality. Their main result is that cheaper credit for poor households increased house prices substantially at the low end of the market. Their model is in the traditional two-sided assignment model framework, where the two sides of the market—buyers and houses—are separate populations, and the prices of other houses do not affect the sellers. In our one-sided model buyers are also sellers, and households' reservation prices as sellers are intricately related to the prices (and qualities available) they face as buyers.

In the theoretical matching literature housing has been used as the motivating example of an indivisible good that needs to be “matched” one-to-one with the buyers. The classic reference is Shapley and Scarf (1974), who present a model where houses are bartered by a discrete set of households who are each endowed with and each wish to consume exactly one house. They show that, regardless of the preference orderings by the households, there always exists at least one equilibrium allocation. Quinzii (1984) and Gale (1984) proved the existence of an equilibrium allocation when, in addition to the indivisible good, there is also a second continuous good or “money.”<sup>2</sup> We get more results by imposing more structure, most importantly we assume homogeneous preferences with housing as a normal good. Also, we assume that there is a continuum of households, which together with the common preference ordering allows us to use calculus for analysis and for an intuitive

---

<sup>2</sup>Quinzii (1984) proved under slightly less general assumptions that equilibrium allocations also coincide with the core.

understanding of comparative statics. With this structure the model can be mapped in a manageable way to the observed joint distribution of house prices and household incomes.

There is a long tradition in explaining heterogeneous land prices in urban economics, going back to the classic Von Thünen model, and Alonso (1964). In these models the exogenous heterogeneity of land is due to distance from the urban center. The focus is on explaining how land use is determined in equilibrium, including phenomena such as parcel size and population density. In modern urban economics<sup>3</sup> there are also some models with income effects. Usually heterogeneity of land is modeled as a transport cost, which is a function of distance from the center, and price differences between locations are practically pinned down by the transport cost function.

Recent models with location choice and endogenous land prices include one by Epple and Sieg (1999), who estimate a structural model where public good provision is endogenous akin to Tiebout (1956). Glazer, Kannianen, and Poutvaara (2008) analyze the effects of income redistribution in a setup where heterogeneous land is owned by absentee landlords. They show that the presence of (uniformly distributed) heterogeneity mitigates the impact of tax competition between jurisdictions because taxation that drives some of the rich to emigrate also leads them to vacate high-quality land, allowing the poor to consume better land than before. Eeckhout, Pinheiro, and Schmidheiny (2013) consider the sorting of worker types across cities, while allowing for within-city sorting in a monocentric city model. In Ortalo-Magné and Prat (2010) household location choice between regions is modeled as part of a larger portfolio problem, where each region has a fixed amount of (infinitely divisible) housing capital. Different regions offer different income processes, so location decisions as well as house prices are affected by hedging considerations.

Most dynamic macroeconomic models with housing assume that housing is a homogeneous malleable good. In any given period, there is then just one unit price for housing. An exception is the property ladder structure used by Ortalo-Magné and Rady (2006) and Ríos-Rull and Sanchez-Marcos (2008), where there are two types of houses: relatively small “flats” and bigger “houses”. For our purposes, a two-type distribution would be far too coarse. Most macro housing literature focuses on the time series aspects of a general level of housing prices, and abstracts away from the cross-sectional complications of the market.

There is also a literature that is concerned with the dispersion of house prices between cities, while abstracting away from heterogeneity within cities. Van Nieuwerburgh and Weill (2010) study house price dispersion across US cities using a dynamic model, where

---

<sup>3</sup>See, for example, the textbook by Fujita (1989).

there is matching by individual ability and regional productivity. Within each city housing is produced with a linear technology, but there is a city-specific resource constraint for the construction of new houses. This causes housing to become relatively more expensive in regions that experience increases in relative productivity. Gyourko, Mayer, and Sinai (2006) have a related model with two locations and heterogeneous preferences for living in one of two possible cities. One of the cities is assumed to be a more attractive “superstar” city in the sense that it has a binding supply constraint for land. An increase in top incomes results in more competition for scarce land, thus leading the price of houses in the superstar city to go up. Moretti (2013) presents a model where changes in relative housing prices between two cities can be affected by productivity (demand for labor) and location preference. He finds that a fifth of the observed increase in college wage premium between 1980 and 2000 was absorbed by higher cost of housing, and that the most plausible cause for this is an increase in demand for high-skill workers in regions that attracted more high-skill workers. By contrast, Diamond (2012) finds that well-being inequality grew more than wage inequality over the same period, once accounting for endogenous change in local amenity levels.

### 3 Model

We model a one-period pure exchange economy, where a unit mass of households consume two goods, housing and a composite good. Preferences are described by a utility function  $u$ , same for all households. Houses come in indivisible units of exogenous quality, and utility depends on the quality of the house, denoted by  $x$ . Every household is endowed with and wishes to consume exactly one house. A household’s endowment of the composite good  $y$  can be interpreted as its income. There are no informational imperfections, or other frictions besides the indivisibility of houses. We assume standard preferences:  $u$  is strictly increasing, differentiable, and strictly quasi-concave.

The aim is to find equilibrium prices for a continuum of house types, given the endowment and the preferences. The aggregate endowment can be described by the joint distribution of households,  $F_{xy}$ , over the consumption space,  $\mathcal{S} = [x_0, x_1] \times \mathbb{R}_+$ . We assume that households are distributed without gaps or atoms over  $[x_0, x_1] \times [\theta_0, \theta_1]$ . Figure 1 depicts the consumption space of this economy, with horizontal dimension representing the quantile in the marginal distribution of house quality  $i = F_x(x)$ , and vertical dimension representing the composite good  $y$  or "consumption" for short.

Consider first the problem of an individual household. Figure 1 shows the budget curve of a household endowed with a bundle  $\{\tilde{x}, \tilde{\theta}\}$ , its shape determined by the price

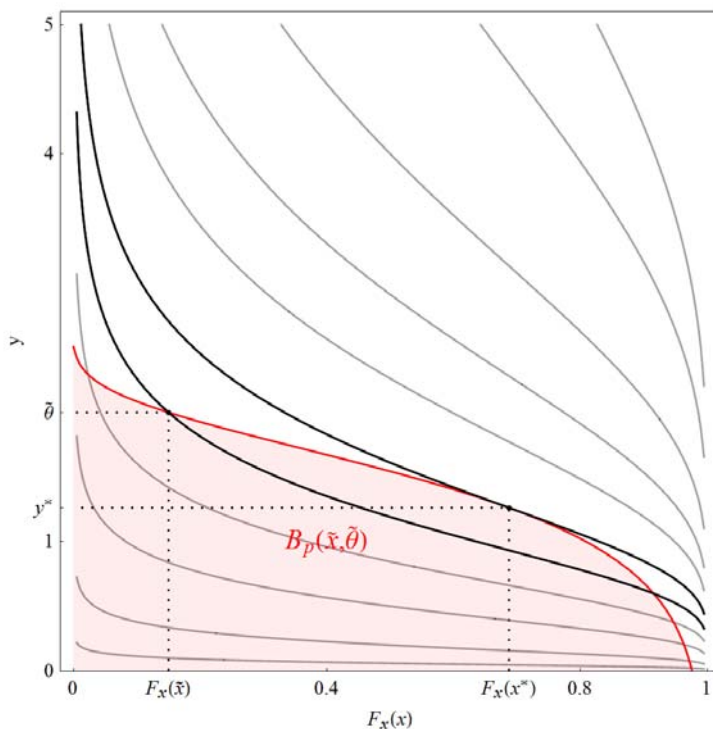


Figure 1: Consumption space. *Horizontal axes depicts the quantile of house quality  $F_x(x)$  and vertical axes the level of non-housing consumption  $y$ . The shaded area is the budget set of a household with endowment  $\{\tilde{x}, \tilde{\theta}\}$  and optimal choice  $\{x^*, y^*\}$ . A set of indifference curves is depicted (for  $u(x, y) = xy$ ); their apparent non-convexity is due to showing the quantile of  $x$  on horizontal axes.*

function  $p$  which households take as given. For this household, the very best houses are outside its budget set, because negative consumption is not feasible. The optimal bundle  $\{x^*, y^*\}$  is where the household gets to the highest indifference curve in its budget set. This household is relatively well endowed in the composite good and would therefore give up some of it in order to trade up to a better house. Formally, the household selects a house of type  $x^* = \arg \max_x u(x, \tilde{\theta} + p(\tilde{x}) - p(x))$ , thus leaving  $y^* = \tilde{\theta} + p(\tilde{x}) - p(x^*)$  for consumption.

In equilibrium, the price function  $p$  must be such that *i*) all households choose their utility-maximizing  $x$  while taking  $p$  as given and *ii*) the resulting allocation is feasible. The indivisibility of houses means that the distribution of house types (the marginal distribution  $F_x$ ) cannot be altered by trading. Thus feasibility requires that, for all  $x$ , the



marginal density of households choosing to live in house type  $x$  is equal to the marginal density of households endowed with type  $x$ .

**Equilibrium conditions** To state the equilibrium conditions more formally, first define the no-trade level of income for a household endowed with a house of type  $x$  and facing price function  $p$

$$y_p^*(x) = \left\{ y \text{ st. } x = \arg \max_{z \in [x_0, x_1]} u(z, y + p(x) - p(z)) \right\}. \quad (1)$$

We later show that, for equilibrium  $p$ ,  $y_p^*$  is single-valued almost everywhere. Now denote the budget set of a household endowed with  $\{\tilde{x}, \tilde{y}\}$  by

$$B_p(\tilde{x}, \tilde{y}) = \{(x, y) \in \mathcal{S} \text{ st. } p(x) + y \leq p(\tilde{x}) + \tilde{y}\}. \quad (2)$$

Finally, the demand and supply of every type of house is equal and all households are maximizing utility under prices  $p$  when

$$\int \int_{B_p(x^*, \max y_p^*(x^*))} f_{xy}(x, y) dx dy = F_x(x^*) \quad (3)$$

for all  $x^* \in [x_0, x_1]$ . (We use  $f$  to denote the density associated with  $F$ .) Then, by Walras' Law, the total consumption of  $y$  also equals its total endowment.

In terms of Figure 1, the resource constraint requires that the proportion of households with an endowment located below the budget curve for the endowment  $\{x^*, y^*\}$  is equal to  $F_x(x^*)$ , the proportion of houses that are of quality  $x^*$  or less. The resource constraints involve double integrals, but, by discretizing the house types, the equilibrium can be solved numerically using standard methods. For most of this paper we will focus on the post-trade allocation, which simplifies the analysis considerably, and allows us to derive closed-form results for the comparative statics of the price distribution. To get there we need two lemmas.

**Lemma 1** *In equilibrium there is positive assortative matching (PAM) by household wealth and house quality.*

That is, in equilibrium, the ranking of households by wealth and by house quality must be the same. The proof is in the Appendix. In short, the diminishing marginal rate of substitution guarantees PAM: of any two households, the wealthier must live in the better house, or else the two could engage in mutually profitable trade. The twist here is that the ordering by wealth is not known beforehand, because the value of the

house is endogenous. So, despite PAM, the equilibrium allocation is not obvious and depends on the shape of the joint distribution of  $F_{xy}$ . The benefit of Lemma 1 is that it guarantees that the equilibrium allocation is essentially one-dimensional, so we can index both households and houses by the house quality quantile  $i$ . We prove the existence of an equilibrium allocation in the Appendix under a finite (but arbitrarily large) number of house types. The equilibrium is associated with unique prices, up to an additive constant that can be interpreted as the opportunity cost of the worst house.

**Lemma 2** *In equilibrium all households  $i \in [0, 1]$  are located on a curve  $\{x(i), y^*(i)\}$  in the consumption space that is continuous almost everywhere. If there are jumps they are upwards.*

This follows directly from Lemma 1: as wealth and therefore utility are increasing in  $i$ , downward jumps in  $y^*$  (as a function of  $i$ ) are ruled out. Similarly, allocations supported over any thick region in consumption space would violate PAM. Only upward jumps in  $y^*$  are not ruled out, but there can only be a countable number of them or else  $y^*$  would not stay finite. Hence  $y^*$  is continuous almost everywhere.<sup>4</sup> However,  $y^*$  does not have to be increasing.<sup>5</sup>

The increasing curve in Figure 2 depicts the equilibrium allocation for a lognormally distributed endowment with positive correlation. Households below the curve are the net suppliers of quality: they are endowed with a relatively high quality house and trade down in order to increase their consumption of the composite good. Households endowed with a house of quality  $x(i)$  and income level  $\theta(i) = y^*(i)$  do not trade. The end points of the equilibrium curve are necessarily  $\{0, \theta_0\}$  and  $\{1, \theta_1\}$ , as the unambiguously poorest and richest households in this economy have either nothing to offer or gain in exchange.

The resource constraints can be illustrated in terms of Figure 2. The mass of households below the budget curve that goes through  $(F_x(x^*), y^*)$  is the mass of households that, at prices  $p$ , consume houses of type  $x^*$  or worse; this is the demand for houses  $x \leq x^*$ . The supply of these types of houses is simply the endowment,  $F_x(x^*)$ , the quantile of house type  $x^*$ . An analogous condition must hold at all  $x^* \in [x_0, x_1]$ . Another way of illustrating the resource constraint is to note that those above the budget curve but to the left of  $F_x(x^*)$  provide the net supply of houses of types below  $x^*$ , and they are trading up. This net supply of house types  $[x_0, x^*]$  must equal the net demand from

---

<sup>4</sup>The jumps are where  $y_p^*$  of (1) could be multi-valued, hence the max in front of it in (3).

<sup>5</sup>For example, a social planner with a taste for equality of utility allocates less  $y$  for those who get the best  $x$  (and it is an equilibrium allocation). For details see the working paper version of this paper.

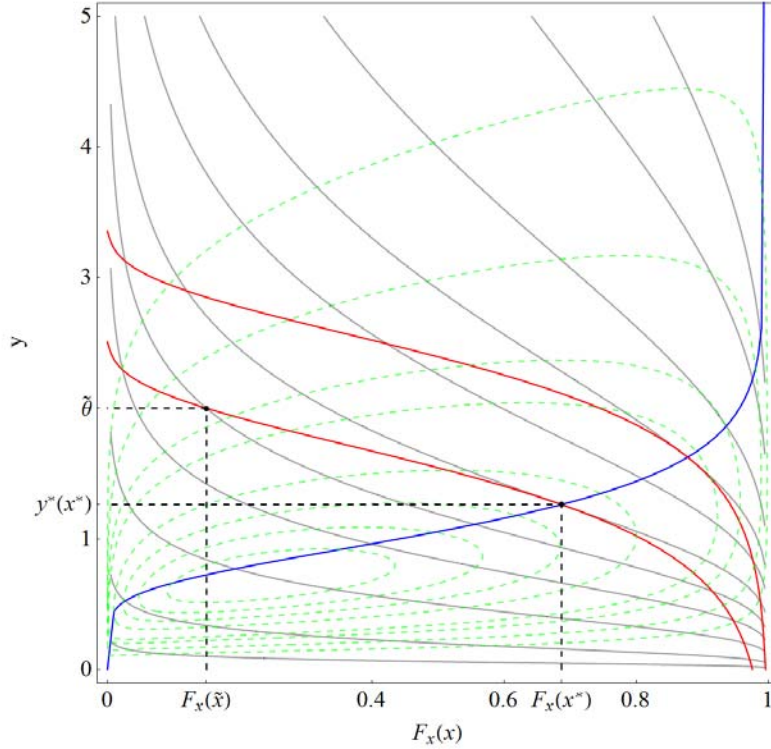


Figure 2: Equilibrium allocation. *The increasing curve shows the equilibrium allocation. Two budget curves are depicted. Dashed curves depict contours of the endowment distribution, i.e., the distribution of households before trade. This example is calculated for a lognormal distribution with  $\mu_x = \mu_y = 0$ ,  $\sigma_x = \sigma_y = 0.5$ ,  $\rho_{xy} = 0.5$ , and utility  $u(x, y) = xy$ .*

those trading down, which is equal to the mass of households to the right of  $F_x(x^*)$  and under the budget curve.

We have now characterized what the allocation must look like after all trading opportunities have been exhausted. From now on we will restrict our analysis to this post-trade world. In a pure exchange economy, the post-trade allocation can be interpreted as just another endowment.

### 3.1 Equilibrium price gradient

Suppose that all trading opportunities have been exhausted, so that the current allocation is an equilibrium allocation. From now on we denote households by the quantile of their house quality,  $i$ . Let's denote by  $\theta(i)$  the allocation of composite good for owners of

houses of quality  $x(i)$ . By definition, at equilibrium prices  $p$  every household prefers to live in its own house, so that

$$i = \arg \max_{j \in [0,1]} u(x(j), \theta(i) + p(i) - p(j)) \quad (4)$$

holds for all  $i \in [0,1]$ . Households are atomistic, so they take  $p$  as given. When the associated first-order condition,  $u_x x' - u_y p' = 0$  is evaluated at the optimal choice  $j = i$  the prices cancel out inside the utility function. (That this optimum is global is guaranteed by Lemma 1.) Solving for  $p'$  we obtain an equation for equilibrium prices:

$$p'(i) = \frac{u_x(x(i), \theta(i))}{u_y(x(i), \theta(i))} x'(i). \quad (5)$$

This price gradient is *the* key equation of our model. Combined with the exogenous boundary condition  $p(0) = p_0$  it can be solved for the equilibrium price function  $p$ . The boundary condition can be interpreted as the opportunity cost for the lowest-quality house, or as the reservation price for the poorest household stemming from some exogenous outside opportunity (such as moving to another region). The fact that  $x$  is a good makes it obvious that  $p$  is strictly increasing, and the continuity of  $u$  and  $x$  implies that  $p$  is continuous.<sup>6</sup>

The intuition behind the price gradient (5) is that the price difference between any neighboring houses in the quality order depends only on how much the relevant households—at that particular quantile of the wealth distribution—are willing to pay for that particular quality difference. This depends on their marginal rate of substitution between house quality and other goods, which in general depends on the level of wealth. The price level at quantile  $i$  is the sum of the outside price  $p_0$  and the integral over all price gradients (5) below  $i$ . This is our next proposition.

**Proposition 3** *Suppose  $\theta$  is an equilibrium allocation. The equilibrium price function is then unique up to an additive constant  $p_0$  and given by*

$$p(i) = p_0 + \int_0^i \frac{u_x(x(j), \theta(j))}{u_y(x(j), \theta(j))} x'(j) dj. \quad (6)$$

Note that the equilibrium price at any quantile  $i$  depends on the distributions of housing quality and income at all quantiles below  $i$ . Hence changes at any part of the price distribution spill upwards but not downwards. Loosely speaking, in terms of a discrete setup, this asymmetry in the direction of price spillovers can be understood by considering the problems faced by the richest and poorest households. If the richest

---

<sup>6</sup>If  $\theta$  has a discontinuity, as is allowed by Lemma 2, then  $p$  has a kink.

household were to get even richer this would have no implication on prices, as it would not make the second richest household willing to pay more for the best house. By contrast, were the poorest household to increase its income slightly (but so that it still remained the poorest), this would increase its willingness to pay for the second worst house, thus increasing the second poorest household's opportunity cost of living in its house. This, in turn, will increase the second poorest household's willingness to pay for the third worst house, and so on, causing the local price increase at the bottom keep spilling upwards in the distribution.

**The case with CES** For the empirical application we assume CES utility,

$$u(x, y) = (\alpha x^\rho + (1 - \alpha) y^\rho)^{\frac{1}{\rho}}, \quad \text{where } \rho < 1 \text{ and } \alpha \in (0, 1), \quad (7)$$

with Cobb-Douglas utility defined in the usual fashion at  $\rho = 0$ . Under the post-trade assumption, where wealth is equal to the sum of income and the equilibrium value of one's house,  $\theta(i) + p(i)$ , the prices cancel out in the right-hand side and this can be solved as

$$p(i) = p_0 + \frac{\alpha}{1 - \alpha} \int_0^i \left( \frac{\theta(s)}{x(s)} \right)^{1-\rho} x'(s) ds. \quad (8)$$

When  $p$  and  $\theta$  are observed, then  $x$  can be solved for under a given elasticity parameter  $\rho$ . The other preference parameter,  $\alpha$ , is absorbed by the units of  $x$  and can then be normalized at, say, one half.

### 3.2 Comparative statics

In this section we analyze the comparative statics of equilibrium prices with respect to changes in income distribution. Here we assume that the economy begins in an equilibrium where there is a strictly positive relation between income  $\theta$  and house quality. For brevity, we call this a "regular" equilibrium.

**Definition.** *A regular equilibrium allocation* is one where there is a strictly monotonic increasing relation between household income and house price.

This, in our model, is equivalent to the case where income and wealth are perfectly rank correlated. Due to PAM (Lemma 1) the equilibrium in our setup has to satisfy perfect rank correlation between wealth and house price, which is a weaker requirement.<sup>7</sup> The purpose of this simplification is to make sure that the analytics of the equilibrium

---

<sup>7</sup>In our empirical application the monotonic relation between income and house prices emerges very naturally as a "side-effect" of kernel smoothing the data under the minimal assumptions needed to make wealth monotonic in house price.

allocation can be used even under changes in income distribution, as order-preserving changes in incomes are then guaranteed to keep the ranking of households by wealth unchanged. A change in the ordering by wealth would generate trading and would thus not fall within the scope of the no-trade case.

It is worth noting that an increase in income inequality does not necessarily lead to an increase in the prices of best houses.

**Proposition 4** *Suppose that the endowments form a regular equilibrium allocation, and that the income distribution experiences a mean-preserving and order-preserving spread where incomes decrease below quantile  $h \in (0, 1)$  and increase above  $h$ . Then housing prices will either i) decrease everywhere or ii) decrease everywhere except at quantiles  $(h', 1]$ , where  $h' > h$ .*

**Proof.** Denote the new distributions by hats. By definition, the new income distribution satisfies

$$\begin{aligned}\hat{\theta}(i) &< \theta(i) && \text{for } i \in [0, h), \\ \hat{\theta}(i) &> \theta(i) && \text{for } i \in (h, 1], \\ \int_0^1 (\hat{\theta}(j) - \theta(j)) \, dj &= 0.\end{aligned}$$

Applying (6), the change in prices at any  $i \in [0, 1]$  is

$$\begin{aligned}\hat{p}(i) - p(i) &= p_0 + \int_0^i \frac{u_x(x(j), \hat{\theta}(j))}{u_y(x(j), \hat{\theta}(j))} x'(j) \, dj - \left( p_0 + \int_0^i \frac{u_x(x(j), \theta(j))}{u_y(x(j), \theta(j))} x'(j) \, dj \right) \\ &= \int_0^i \left( \frac{u_x(x(j), \hat{\theta}(j))}{u_y(x(j), \hat{\theta}(j))} - \frac{u_x(x(j), \theta(j))}{u_y(x(j), \theta(j))} \right) x'(j) \, dj.\end{aligned}\tag{9}$$

The inverse of the marginal rate of substitution between  $y$  and  $x$ ,  $u_x(x, \theta)/u_y(x, \theta)$ , is increasing in  $\theta$ , and  $x' > 0$ , so the integrand in (9) is negative at all  $j$  where  $\hat{\theta}(j) < \theta(j)$ , i.e., for  $j < h$ . Similarly, the integrand is positive for  $j > h$ . The definite integral in (9) must therefore be strictly negative at  $i = h$ , where it reaches its minimum, and increasing above  $h$ . If, at some  $h' > h$  the definite integral reaches zero then it will be positive at all  $i > h'$ , but it might not reach zero before  $i = 1$ , in which case the price change is negative at all  $i \in (0, 1]$ . ■

Intuitively, if income is redistributed from poor to rich, this will increase the local price gradient (5) at the top quantiles, as the willingness-to-pay for extra quality goes up for the rich. But, for the same reason, the price gradient at bottom quantiles goes down. Above  $h$ , the change in the price gradient is positive, but the negative spillover from below

will dominate until some  $h' > h$ , and it can be that the cumulative impact of positive gradients is not enough to overtake the negative impact. It is therefore possible for all house prices to go down in response to an increase in inequality. This would happen, for instance, if the quality differences between best houses are relatively small so that the rich are less inclined to use income growth to bid up the price of the next better house.

The impact of a simple increase in income levels is characterized by the next Proposition.

**Proposition 5** *Suppose the endowments form a regular equilibrium allocation. If incomes rise at all quantiles  $i \in [i_0, i_1)$  in an order-preserving manner, then housing prices will increase at all quantiles  $i \in (i_0, 1]$ , and this increase is increasing in  $i$ .*

The reasoning is similar as in the Proof of Proposition 4, but even simpler because the new price gradient is greater than the original gradient at all quantiles. In the simplest special case all incomes rise in an order-preserving manner; then housing prices rise at all quantiles  $i \in (0, 1]$ , and the more so at higher  $i$ .

As an immediate corollary, an increase in income levels will also increase the variance of house prices. A simple extreme case is where all incomes catch up with the highest income level. Proposition 4 tells us that such a complete elimination of income inequality would increase both the levels and the dispersion of house prices. Intuitively, all households would then be competing for the best houses and the price difference between a low quality and a high quality house must become relatively large to make some household willing to hold the low quality house.

The exogenous bottom price  $p(0)$  is held constant under comparative statics, so the price changes we are referring to are for quantiles  $(0, 1]$ . A change in the lowest price,  $p_0$ , would cause all prices to change by that same amount. As the model does not include the possibility to move out of the market, the level of the constant is irrelevant for the attractiveness of any trade: it increases both the buying and selling prices, and thus washes out of all transactions.

### 3.3 Fixed outside opportunity and endogenous city size

Now let's consider the possibility that the city size (population) can change in response to changes in distributions. For this purpose let the quantile  $i$  in  $x(i)$  and  $\theta(i)$  denote the quantile in the potential population, and use  $m$  to denote the quantile of the lowest types in the actual population. In equilibrium, the marginal household  $m$  is indifferent between living in the model city and the outside opportunity. As before, the opportunity

cost of a house in the city is fixed at  $p_0$ , which could be, for example, the value of land in the next best alternative use. The basic model with a fixed city size can be interpreted as a renormalization of the population size  $1 - m$  to unity, where the quantile  $m$  in the potential population is then located at quantile 0 in the actual population.

Now suppose that a fixed utility  $v_0$  is available outside for anyone who chooses to not live in the model city. This can be interpreted as there being a large outside economy that is not affected by what happens in the model city. The marginal household is determined by

$$u(x(m), \theta(m)) = v_0. \quad (10)$$

Denote the solution as  $m_0$ . It does not depend on  $p_0$ , because outside utility does not depend on wealth.

Consider the impact of an increase in income inequality, as we did in Section 3.2 for a fixed city size. Now inequality must be thought to affect the distribution for the potential population, not just for those who live in the city. The decrease in income causes the utility level inside the city to be decreased for the marginal household  $m_0$ , so it moves out of the city.<sup>8</sup> The city size contracts, and the new margin is further up the distribution,  $m' > m_0$ . The vacated houses sell at the opportunity cost  $p_0$ , and are no longer part of the price distribution. For households that remain in the city the price gradient (5) is exactly the same as with a fixed city size, so the same comparative statics (Proposition 4) still apply. The difference is that the integral that defines the price level (6) "skips" the bottom segment that moved out  $[m_0, m']$ , so all prices above  $m'$  are lower by a constant  $p(m') - p_0$  than they would be under a fixed city size. The bottom line is that having a fixed outside opportunity, as opposed to a closed city, means that an increase in income inequality will cause the city to contract in size, and will lower house prices, but does not affect the price differences between occupied houses.

**Alternative formulation: fixed outside house-price option** Now suppose that instead of a fixed utility level, the outside opportunity consists of a fixed quality level of houses  $\tilde{x}$  is at a fixed price  $\tilde{p}$ . This can be interpreted as an option to construct a new house of quality  $\tilde{x}$  at cost  $\tilde{p}$  somewhere outside the model city. The marginal household is determined by

$$u(x(m), \theta(m)) = u(\tilde{x}, \theta(m) + p_0 - \tilde{p}). \quad (11)$$

---

<sup>8</sup>We assume here that the margin is clearly below the mean of the potential distribution, so that an increase in inequality reduces the income of the marginal household.



Consider an increase in inequality of  $\theta$ . Now the change in income affects also the utility from the outside opportunity. To see what happens to city size we need to pin down what happens to the utility difference  $u(x_0, \theta) - u(\tilde{x}, \theta + p_0 - \tilde{p})$  when  $\theta$  is decreased. The direction of change for city size depends on the "quality type" of the city. If this is a city of relatively high quality, and thus of high price, at the margin ( $x(m_0) > \tilde{x}, p_0 > \tilde{p}$ ), then income reduction will make it less attractive to its poorest households and it will contract at the margin. Conversely, if the model city is of relatively low quality at the margin then the outside opportunity becomes less attractive to the marginal household, and the increase in inequality causes the city to expand.

### 3.4 Two-Sided Matching

In traditional assignment models the matching is two-sided, which means that there are two separate populations that take up the different sides of a match. The two-sided setup is simplest to interpret as a rental market. In urban economics models with heterogeneous land it is standard to assume that all land is initially owned by competitive outside sellers or "absentee landlords." Again, the crucial simplification is that sellers' reservation prices are exogenous (and not binding). In a two-sided model buyer wealth is exogenous (and seller wealth irrelevant) so it lacks "the wealth channel" of the one-sided setup. The two-sided version of our model is essentially the same setup as was analyzed by Braid (1981), who showed the equivalent of our Proposition 5 for the rental market.

Consider now the buyer (or renter) at quantile  $i$  of the income distribution under two-sided matching. By Lemma 1, equilibrium must involve positive assortative matching by buyer wealth (now just  $\theta$ ), and house quality. Thus  $p$  must result in every household buying a house of the same quality rank as is their rank in the income distribution.

$$i = \arg \max_{j \in [0,1]} u(x(j), \theta(i) - p(j)) \quad \text{for all } i \in [0, 1] \quad (12)$$

Now the price of the house chosen is not part of household wealth and so it does not cancel out of the price gradient. Equilibrium prices are defined as a nonlinear ordinary differential equation:

$$p'(i) = \frac{u_x(x(i), \theta(i) - p(i))}{u_y(x(i), \theta(i) - p(i))} x'(i). \quad (13)$$

Combined with a boundary condition this can still be solved for the equilibrium price function  $p$ . The boundary condition must now satisfy  $p_0 \leq \theta(0)$ , or else the poorest household cannot afford to live anywhere.

Under CES utility (7) the equilibrium price gradient (13) is

$$p'(i) = \frac{\alpha}{1 - \alpha} \left( \frac{\theta(i) - p(i)}{x(i)} \right)^{1-\rho} x'(i). \quad (14)$$

## 4 Inferring quality and computing counterfactuals

For the equilibrium price relation to be applied to cross-sectional data it is necessary to assume that observed prices correspond to equilibrium prices that emerge after all trading opportunities have been exhausted. We think this is a reasonable interpretation of data because, at any point in time, only a small fraction of households trade houses. Under this assumption the unobserved distribution of house qualities can be inferred from the observed relation between household income and house prices. That is, for given distributions  $p$  and  $\theta$ , and for a given utility function  $u$ , the distribution of  $x$  can be inferred. (We discuss the inference of preferences in the next section.) This can be done by treating  $x$  as the unknown in the differential equation (5), while normalizing the boundary condition  $x(0) = x_0$  at any positive value.

Having inferred the distribution of  $x$  based on the observed distributions of  $\theta$  and  $p$ , any counterfactual income distribution  $\hat{\theta}$  then pins down a counterfactual distribution of house prices, by combining  $\hat{\theta}$  and  $x$  in the (possibly discrete equivalent of) equilibrium price relation (8). Note, however, that as  $p_0$  is exogenous, our model only explains the differences in prices relative to the marginal unit of housing,  $p - p_0$ . In our counterfactuals the lowest price is always taken to be the lowest price in the data.

Inference requires data on prices (house values)  $p_0 < p_1 < \dots < p_N$  and associated household incomes  $\theta_h$ . The minimal requirement for the data to be consistent with the model is that observed wealth  $p + \theta$  must be in strictly positive relation with  $p$  (Lemma 1). This relation is, of course, not perfect in reality, but it emerges very naturally in our data with kernel smoothing (more of which in Section 5.1).

The basis for inference is the "incentive compatibility" condition that makes household  $h$  want to buy its equilibrium match—which is house  $h$ —instead of any other house  $h'$ . For CES utility, these conditions are

$$(\alpha x_h^\rho + (1 - \alpha) (\theta_h)^\rho)^{\frac{1}{\rho}} \geq (\alpha x_{h'}^\rho + (1 - \alpha) (\theta_h + p_h - p_{h'})^\rho)^{\frac{1}{\rho}}, \text{ for all } h, h' \in \{0, \dots, N\}. \quad (15)$$

Thanks to PAM, we know that this constraint is only binding for household pairs that are "neighbors" in the rank by house quality ( $h' = h - 1$ ). For convenience, we assume that these constraints hold as an equality, and we obtain a discrete equivalent of the price

gradient.<sup>9</sup>

$$p_h = p_{h-1} + \left( (\theta_h)^\rho + \frac{\alpha}{1-\alpha} (x_h^\rho - x_{h-1}^\rho) \right)^{\frac{1}{\rho}} - \theta_h \quad (16)$$

Equivalently, solving for  $x_h$  yields the inference formula

$$x_h = \left( x_{h-1}^\rho + \frac{1-\alpha}{\alpha} [(\theta_h + p_h - p_{h-1})^\rho - (\theta_h)^\rho] \right)^{\frac{1}{\rho}}. \quad (17)$$

Note that the value of  $\alpha \in (0, 1)$  is without consequence when  $x$  is unobserved, because changing  $\alpha$  is equivalent to changing the units of  $x$ . We set  $\alpha = 1/2$ . (The levels of  $x$  are not themselves meaningful, and this normalization is without loss of generality for the counterfactual prices). Denoting  $\hat{x} = x^\rho$ , (17) can now be solved as

$$\hat{x}_h = \hat{x}_0 + \sum_{j=1}^h [(\theta_j + p_j - p_{j-1})^\rho - (\theta_j)^\rho], \quad (18)$$

which includes an undefined constant of integration  $\hat{x}_0$ , i.e. the quality of the worst occupied house.

Finally, by combining inferred  $\hat{x}_h$  with observed  $p_0$  and with posited counterfactual  $\hat{\theta}_h$  in the price difference equation (16), we obtain counterfactual prices  $\hat{p}_h$ . In the CES case the steps can be combined to yield the counterfactual prices directly as

$$\hat{p}_h = \hat{p}_{h-1} + \left( (\theta_h + p_h - p_{h-1})^\rho - (\theta_h)^\rho + (\hat{\theta}_h)^\rho \right)^{\frac{1}{\rho}} - \hat{\theta}_h \quad (19)$$

for  $h = 1, 2, \dots$ , with  $\hat{p}_0 = p_0$  again taken as exogenous.

## 5 Empirical application: income inequality and house prices

In this section, we apply the model to data. Our aim is to further illustrate the mechanisms of the model and to evaluate their quantitative importance. We first present the data and describe the calibration. We then show how the income elasticity of housing demand depends on the distributions of quality and income, and how it varies between households. In our main counterfactual exercise, we use the model to estimate the impact of the recent increase in income inequality on house prices. We consider a counterfactual income

---

<sup>9</sup>In the discrete setup there is a match-specific rent that the "neighbors" could bargain over. Here we in effect assume that "sellers" have all bargaining power. Making the opposite assumption makes practically no difference to our empirical results.

distribution for 2007 where all incomes grow uniformly since 1998 at the same rate as the actual mean income. We compare house prices that result from the counterfactual income distributions with the prices that result from the actual income distribution. We then repeat the analysis under two alternatives formulations. We compare these baseline results with those obtained using a two-sided version of our model (see Section 3.4), where wealth is exogenous. This serves to gauge the importance of the wealth channel in our model. Finally, we also compare our results with a setup where housing capital is perfectly divisible.

We abstract from changes in other house price determinants and focus entirely on changes in income distribution. Of course, this is not to say that changes in income distribution are the dominant driver of house price changes. Rather, our main counterfactual exercise can be seen as a way to disentangle the impact of increased income inequality on house prices from changes in various other house price determinants.

Our model abstracts from many real world features that may influence the relationship between the distributions of income and housing prices. In particular, as discussed in section 3.3, migration may be responsive to changes in income distribution. We do not attempt to account for migration. That is, we hold the population constant in our counterfactuals. We also don't allow for adjustments to housing structures and keep the quality distribution constant. In reality, the value of the house consists of the value of land and the value of structures and it is possible to adjust the quality of the structures to some extent. Even in the long term, according to Davis and Heathcote (2007), most growth and volatility in U.S. house prices has been due to variation in land values.<sup>10</sup>

## 5.1 Data

We use income and price data from the American Housing Survey (AHS) for six metropolitan areas (MA, or "city"): Baltimore, Boston, Houston, Minneapolis, Tampa, and Washington. Our income measure is total disposable income, including taxes and transfers, during the last year. "Houses" include homes in multi-unit dwellings. The choice of MAs was determined by the availability of data. For a more detailed description of the data, see Appendix B.

House prices in the AHS data have been censored separately at each MA at the 97% percentile. Thus we have to exclude the top 3% from our analysis. Furthermore, there are apparently significant data quality issues at the bottom of the price distributions, with

---

<sup>10</sup>Davis and Palumbo (2008) estimate the share of land in house values for detached single-family homes; in our sample cities it ranges between 31% (Houston) and 76% (Boston) in 2004.

many house prices observed in the range of a few hundred or thousands of dollars. For this reason we drop the bottom 5% of houses in each MA, and so the 5th percentile price will be the lowest (and thus exogenous) house price in our analysis. All of our results refer to this restricted sample.

We observe the joint distribution of income  $\theta$  and house price  $p$  in each of the six MAs for both 1998 and 2007. There is a strong positive relationship between house value and income. Rank correlation between income and house price is between 0.38 and 0.52 depending on MA (see Table B1 in the Appendix). To be consistent with the equilibrium of our model, the levels of wealth  $w = \theta + p$  should be perfectly rank correlated with house value  $p$  (recall Lemma 1). Thus we reduce the relation of income and house price into a curve, by using kernel regression to estimate  $\theta(i)$  as  $E[\theta|F_p(p) = i]$ , where  $F_p$  is the empirical CDF of  $p$ .<sup>11</sup>

Figure 3 displays the distributions of smoothed income relative to its mean in 1998 and 2007. The only apparently exceptional case is Houston where lowest incomes have grown faster than the mean. Table B1 displays Gini coefficients for our estimated permanent income (i.e. smoothed income), annual income, and house value. All MAs feature an increase in income inequality from 1998 to 2007. Housing values show no such systematic pattern.

We need to set a value for the interest rate  $r$  to make the units of yearly income compatible with the house price: income  $\theta$  is measured as annual income divided by the interest rate. (In effect, we assume that households face an infinite time-horizon, and expect no changes in the future.) We set  $r = 0.05$ ; changing  $r$  over a reasonable range (2 – 8%) makes little difference to our inferred elasticity parameter  $\rho$  or the results of the counterfactual experiments that we present below.

## 5.2 Inferring preferences

To apply the model we need a value for the preference parameter  $\rho$  in the CES-utility function (7).<sup>12</sup> This parameter determines the elasticity of substitution between consumption and housing,  $1/(1 - \rho)$ . We can estimate  $\rho$  using the following procedure. First, for a given  $\rho$  and for the observed  $\{\theta^{98}, p^{98}\}$ , we infer the quality distribution in 1998, de-

---

<sup>11</sup>We use the Epanechnikov kernel and a bandwidth of 9%, except in Tampa where a bandwidth of 11% is required for the smoothed data to conform to assortative matching by wealth and house price.

<sup>12</sup>Li, Liu and Yau (2009) and Bajari, Chan, Krueger and Miller (2010) use a life cycle model to estimate the same preference parameter. In contrast to these papers, we exploit changes in the cross-sectional distribution of housing prices. This is possible in our model because housing prices are in general a non-linear function of housing quality.

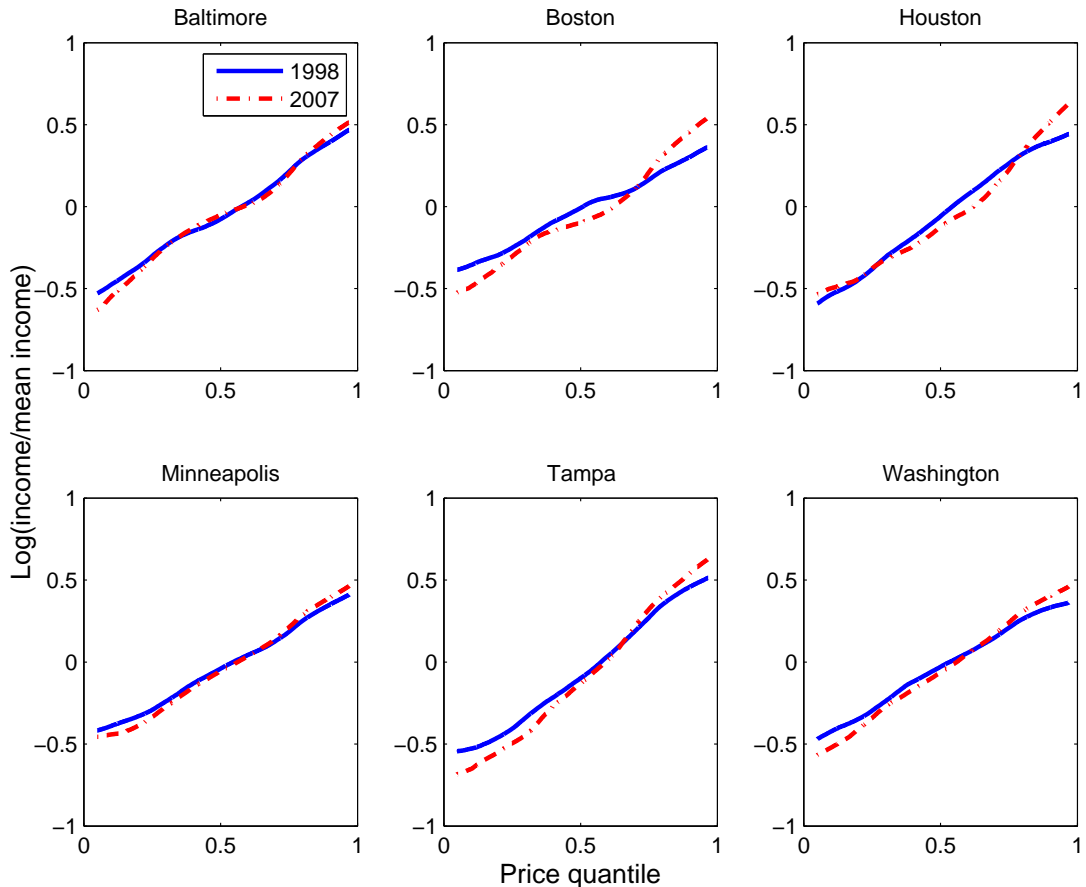


Figure 3: *Distribution of permanent income relative to mean by metropolitan area.*

noted by  $x^{98}$ , using the inference formula in (18). Then we then use the equilibrium price formula—the discretized equivalent of (8)—to predict the 2007 housing price distribution given  $x^{98}$  and the observed wealth in 2007 ( $\theta^{07} + p^{07}$ ). In other words, assuming that the quality distribution did not change over time, we ask what would be the predicted price distribution in 2007 given the observed 2007 income distribution. We denote this prediction by  $\hat{p}^{07}$ . We have to implicitly assume that the ranking of households by income has not changed, as we apply a static model in both years. We repeat this for a range of values for  $\rho$ , and pick the value that gives the best prediction for the 2007 price distribution.<sup>13</sup>

This inference procedure relies on the assumption that the distribution of income is the only determinant of house prices that has changed. Changes in other determinants

---

<sup>13</sup>We minimize the mean of absolute percentage errors (MAPE), i.e. the mean value of  $|\log(p^{07}(i)) - \log(\hat{p}^{07}(i|\rho))|$ .

could bias our estimate of the elasticity parameter. Probably the most important issue in this respect is migration and new construction. In order to mitigate this problem, we choose our baseline value for the elasticity parameter based on data from the three MAs (of the six in our data) where population has grown the least from 2000 to 2010 according to the US Census. These MAs are Baltimore (6% population growth from 2000 to 2010), Boston (4%), and Minneapolis (11%). The inferred values for the elasticity of substitution are 0.42 (Baltimore), 0.63 (Boston) and 0.48 (Minneapolis). For comparability, we use the same elasticity in all MAs. Based on these results, we set the elasticity of substitution in the benchmark case equal to 0.5 (i.e.,  $\rho = -1$ ). We also consider elasticities equal to 0.4 and 0.6 in our main counterfactual exercise below, in order to measure the importance of the elasticity parameter. Table B2 in the appendix shows the inferred elasticity of substitution for all MAs and figure B1 the relative prediction errors,  $\log(\hat{p}^{07}/p^{07})$ , by quantile of house value.

### 5.3 Income elasticity of housing expenditure

Our inferred elasticities are in line with many studies that use household data.<sup>14</sup> However, as always in structural estimation, the interpretation of the parameters depends on the specifics of the model. Our estimates are not directly comparable with those obtained in studies that do not take into account the friction arising from the indivisibility of houses. In our setup, the effective income and price elasticities of housing demand are not determined solely by assumptions about preferences, because prices are a nonlinear function of quality. This nonlinearity also implies that income and price elasticities of housing expenditure will vary across income levels (despite CES utility).

In order to illustrate these features, let us define income elasticity of housing demand at quantile  $i$  using a counterfactual increase in housing expenditure that would result if a single household were to alone experience a change in income. Following a  $\Delta$  percent change in income at quantile  $i$ , the household will reoptimize its consumption. Housing expenditure changes from  $p(i)$  to  $p(j)$ , where

$$j = \arg \max_{s \in [0,1]} (u(x(s)), (1 + \Delta)\theta(i) + p(i) - p(s)) \quad (20)$$

is the household's new quantile in the distributions of wealth and housing quality, given that everyone else's incomes stay the same. In equilibrium, this elasticity can be determined just from current prices and incomes. Since there is positive assortative matching by housing quality and wealth, we can find the new housing expenditure  $p(j)$  simply by

---

<sup>14</sup>See e.g. Li, Liu, and Yao 2009, and the references therein.

finding where the household will be located in the wealth distribution after its income is changed. Thus  $j$  from (20) solves  $(1 + \Delta)\theta(i) + p(i) = \theta(j) + p(j)$ . The change in wealth position is independent of the utility function (as long as it exhibits the diminishing MRS required for positive sorting). In the end, this result stems from the assumption that houses are indivisible and have fixed qualities.

Figure 4 shows the income elasticity of housing expenditure in each MA. For each quantile  $i$ , the elasticity is computed as the midpoint arc elasticity around  $\theta^{07}(i)$  with a 10% income change.<sup>15</sup> We can calculate this elasticity only up to the point where the hypothetical increase in income that would lift the households above the top-coding threshold. Figure 4 reveals how income elasticity varies quite substantially over the distribution. Intuitively, how much more one would spend on housing following an increase in income depends on the increase in housing quality that would be available. The additional housing quality that one extra dollar can buy varies over the distribution, because the additional housing quality that a dollar can buy varies over the distribution, as it depends on the shape of the quality distribution and on the incomes of other buyers and sellers.

The observation that the elasticity of demand varies substantially over the distribution suggest that estimating household preferences without taking into account the indivisibility and heterogeneity of houses is problematic. For instance, the effects of a given change in aggregate income on aggregate housing demand can depend on the associated changes in the distribution of income. This is also illustrated by the counterfactual experiments in the next section.

The assessment of individual household's demand elasticity above involved a counterfactual where prices are held constant because one household has a vanishingly small impact on prices. If, by contrast, all households experience a change in income by the same percentage, then standard intuition about elasticity is left intact. Let's consider the price impact of all households simultaneously receiving more income. When all household incomes go up by  $\Delta$  percent then prices must react; by how much depends on preferences. (Now no one will actually move, because everyone's rank in wealth distribution stays the same.) By substituting into (8), we see that, under CES utility, all prices of inframarginal houses (price differences  $p - p_0$ ) blow up by a factor of  $(1 + \Delta)^{1-\rho}$ .

---

<sup>15</sup>Using a smaller  $\Delta$  results in otherwise similar but more "erratic" elasticity curves. This is because the price distribution is not smoothed, so small changes in  $i$  can result in large changes in  $p(i)$ .



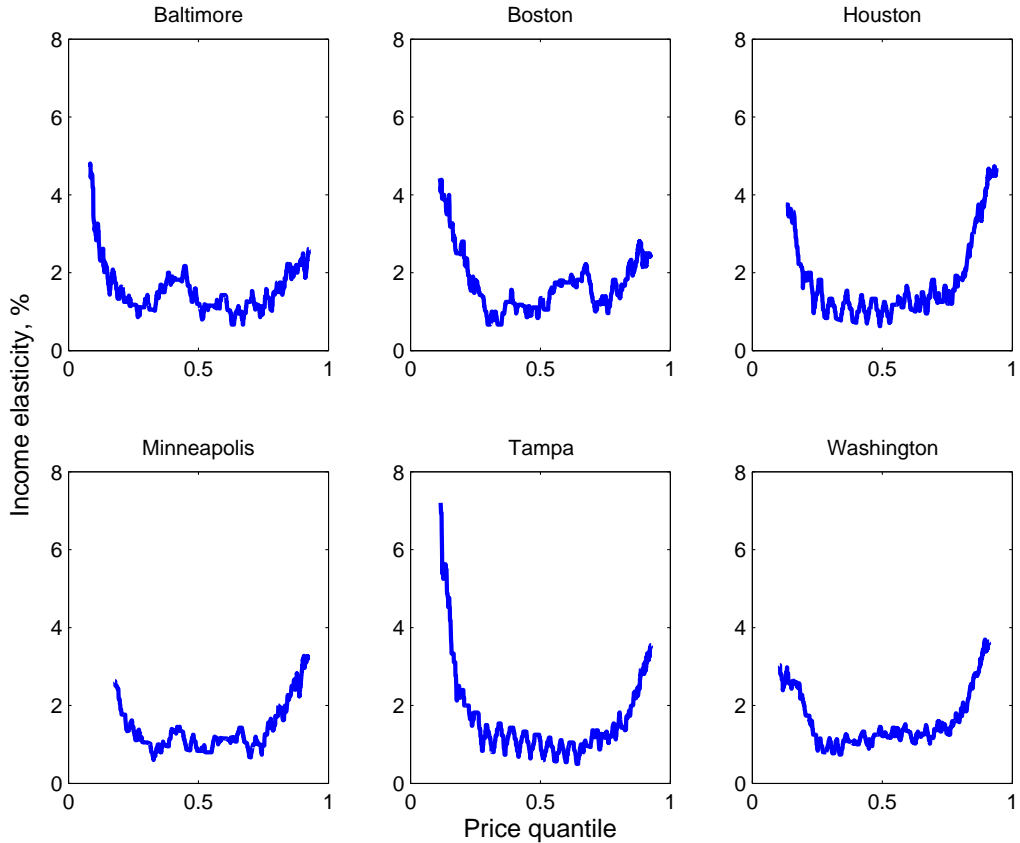


Figure 4: *Estimated income elasticity of housing expenditure by wealth quantile.*

## 5.4 Counterfactuals

We construct counterfactuals to assess the impact of changes in the distribution of income between 1998 and 2007 on housing prices in six US metropolitan areas. Specifically, given  $\rho$ , we infer the quality distribution  $x^{98}$  in each MA and compute the predicted price distributions in 2007 under a counterfactual income distribution that has the same shape as in 1998 but the same mean as in 2007. We then compare this counterfactual price distribution with the fitted price distribution that is obtained by plugging in the actual 2007 income distribution to the model; the difference between the two is our estimate of the contribution of increased inequality on house prices.

Figure 5 displays the relative difference between counterfactual and empirical housing prices. Consider first the benchmark case. The impact of increased income inequality is qualitatively similar in all but one MA: it has lowered housing prices until about 80-90th percentile and increased prices in the upper tail. The exception is Houston, where

changes near the bottom of the income distribution work to increase the prices of the lowest quality houses. This reflects the fact, visible in Figure 3, that in Houston, unlike in the other MAs, the lowest incomes have increased relative to the mean income, even though the Gini coefficient has increased there as in other MAs.

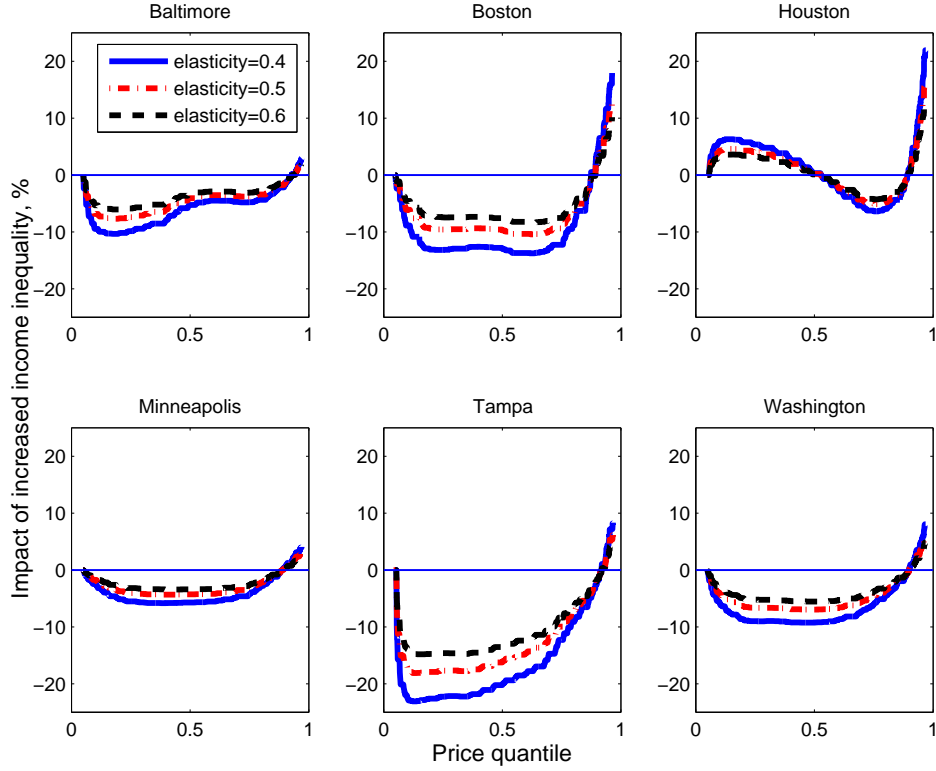


Figure 5: *Impact of increased income inequality on house prices.*

It is a general feature of the results that the lower is the assumed elasticity, the larger is the impact of income inequality on house prices at any given quantile. However, in other respects, the results are not much affected by the assumed elasticity. In particular, the point at which the price effect turns from negative to positive is not much affected by the assumed elasticity. Also, the increase in the price of even the best houses (those at the 97th percentile) is always moderate. Intuitively, the increase in income inequality results in lower incomes at the bottom of the distribution relative to the counterfactual of uniform income growth. This works to lower the prices of the lowest quality houses. We showed in Section 3.2 how this negative price impact at the bottom spills upwards in the quality distribution. This effect counteracts the local increase in willingness-to-pay among the lower part of the better-off households who in actuality saw their incomes rise

faster than the mean rate of growth.

The impact on average prices is shown in Table 1. The first three columns display the mean relative effect of the change in the shape of the income distribution on house prices for the three elasticities of substitution considered. The last column displays the absolute price change in the benchmark case. The mean impact of change in income inequality in the benchmark specification varies from  $-9.9\%$  in Tampa to  $0.6\%$  in Houston.<sup>16</sup>

Elasticity $1/(1 - \rho)$	Change (%)			Change ( $\$k$ )
	0.4	0.5	0.6	0.5
Baltimore	-4.6	-3.6	-2.9	-11.0
Boston	-7.5	-5.9	-4.7	-34.3
Houston	+0.1	+0.6	+0.4	+1.2
Minneapolis	-3.3	-2.5	-2.0	-7.1
Tampa	-12.1	-9.9	-8.3	-26.2
Washington	-5.2	-4.0	-3.2	-16.9

**Table 1.** Change: The impact of the change in the shape of the income distribution on house prices between 2007 and 1998, relative to what house prices would have been under uniform income growth.

## 5.5 Quantifying the wealth channel and the role of indivisibility

In our one-sided assignment model, the distribution of wealth is endogenous as it depends on house prices. As discussed in section 3.4, in a traditional two-sided model, in contrast, the wealth distribution is taken as given. In this subsection we evaluate the quantitative importance of an endogenous wealth distribution. To this end, we compare the impact of increased income inequality in our one-sided model to its impact in a two-sided model.

Recall that under two-sided matching the "buyers" and "sellers" come from separate populations, and there is no feedback from prices to the willingness to buy or sell. In terms of our counterfactual, with owner-occupied housing, this means that everyone in the model city has decided to sell their house, and the value of living somewhere else is not the binding constraint that would determine their reservation price for selling their current house. The sellers' problem is simply to sell the current house at the highest possible price,

<sup>16</sup>In a working paper version (Määttänen and Terviö 2009) we used data from the Helsinki metro region in 1998 and 2004, where there was also a significant increase in income inequality. The results are similar, with a mean effect of  $-3.2\%$ .

so the initial wealth distribution of sellers is inconsequential. Buyer wealth matters but is exogenous, and it includes any possible proceeds from selling a house somewhere else.

In order to make the two-sided model fully comparable with the one-sided model, we use for each city the same quality distribution  $x^{98}$  that was inferred in the main exercise, and then supplement buyers incomes with transfers that are equal to the initial house prices  $p^{98}$ . The point is to give the buyers in the two-sided version the same initial wealth distribution that households have in the one-sided setup. For a given income distribution  $\theta$ , equilibrium house prices  $p$  are then again determined so that buyers are choosing their optimal house while taking prices as given:

$$i = \arg \max_{j \in [0,1]} u(x^{98}(j), \theta(i) + p^{98}(i) - p(j)) \quad \text{for all } i \in [0, 1]. \quad (21)$$

This results in  $p(i) = p^{98}(i)$  for all  $i$ . That is, given the same quality distribution, the empirical 1998 price distribution is then consistent with the 1998 income distribution in both models.<sup>17</sup> This is done by assuming that the value of the houses owned by the buyers (somewhere outside the model city) is exactly the same as is the initial value of houses in the model city before the change in income distribution.

We then evaluate the impact of increased income inequality in the same way as before. We compute the equilibrium price distribution that results from setting  $\theta$  in (21) equal to the actual (smoothed) income distribution in 2007 and compare that with the equilibrium price distribution that result from setting  $\theta$  equal to a counterfactual income distribution that has the same shape as in 1998 but the same mean as in 2007. (However, we assume here, in both models, that  $p(0)$  remains fixed at  $p^{98}(0)$ . Otherwise, the non-housing consumption of the poorest household would change in the one-sided model.) As before, we assume an elasticity of substitution equal to 0.5.

At the same time, we also compare our main results with those resulting from a model with complete divisibility of houses. For this purpose we use a model with perfectly malleable "housing capital", denoted by  $h$ , measured in units of the composite good (non-housing consumption). The equivalent of the house price distribution is now simply the housing capital distribution. Given initial income  $\theta^{98}(i)$ , the initial housing capital  $h^{98}(i)$  can be solved from the household optimization problem:

$$h^{98}(i) = \arg \max_{\tilde{h}} u(\tilde{h}, \theta^{98}(i) + h^{98}(i) - \tilde{h}).$$

This equation is the equivalent of (4). However, the resulting price distribution no longer matches the actual distribution. Predicted and counterfactual housing capital holdings

---

<sup>17</sup>To see this, compare (21) with (4), and take into account how  $x^{98}$  is determined given the 1998 income distribution  $\theta^{98}$  and  $p(0) = p^{98}(0)$ .

$h(i)$  can now be solved from the household maximization problem

$$h(i) = \arg \max_{\tilde{h}} u(\tilde{h}, \theta(i) + h^{98}(i) - \tilde{h}),$$

where  $\theta(i)$  is taken either from the actual 2007 income distribution or from the counterfactual distribution. As there is now only one price in this economy—the relative price of non-housing consumption and housing capital—and it is fixed, now CES utility implies that households spend a constant fraction of their resources on housing. Hence, we can set  $\rho = 0$  without loss of generality for this exercise (i.e., assume Cobb-Douglas utility). On the other hand, the housing share parameter  $\alpha$  in (7) is no longer absorbed by the units of housing quality, and we set it so as to match the ratio of average house value to average income in the data.<sup>18</sup>

Figure 6 displays the impact of increased income inequality on house prices (“one-sided” and “two-sided”) or housing capital (“malleable”) in the three models. The impact of increased income inequality is qualitatively similar in both the one-sided and two-sided assignment models. However, the wealth channel works to magnify the impact as changes in house prices feed back into housing demand via household wealth. In some cities the difference is substantial, but qualitatively the changes are similar across distributions and quantiles. By contrast, the model with malleable housing capital results in a very different distributional impact than the assignment models. This is because, without indivisibilities, a change in housing wealth at any point of the distribution reflects a change in income at that point only, whereas both assignment models feature a form of cumulative spillovers due to households competing for specific houses.

## 6 Conclusion

We have presented a new framework for studying the relationship between the income distribution and the housing price distribution. In the model houses are heterogeneous and indivisible. The key element is that all houses are owned by the households, and, due to concave utility, their reservation prices as sellers depend on the opportunities available to them as buyers, which in turn depend on their incomes and on the prices of other houses. Thus, our model provides a framework for analyzing how income differences get capitalized into house prices. The equilibrium is tractable under the assumption that households prefer to live in their current house.

---

<sup>18</sup>The ratio varies across the MAs. We match the non-weighted average ratio which is 19%. This ratio is computed after annual income has been divided by the interest rate.

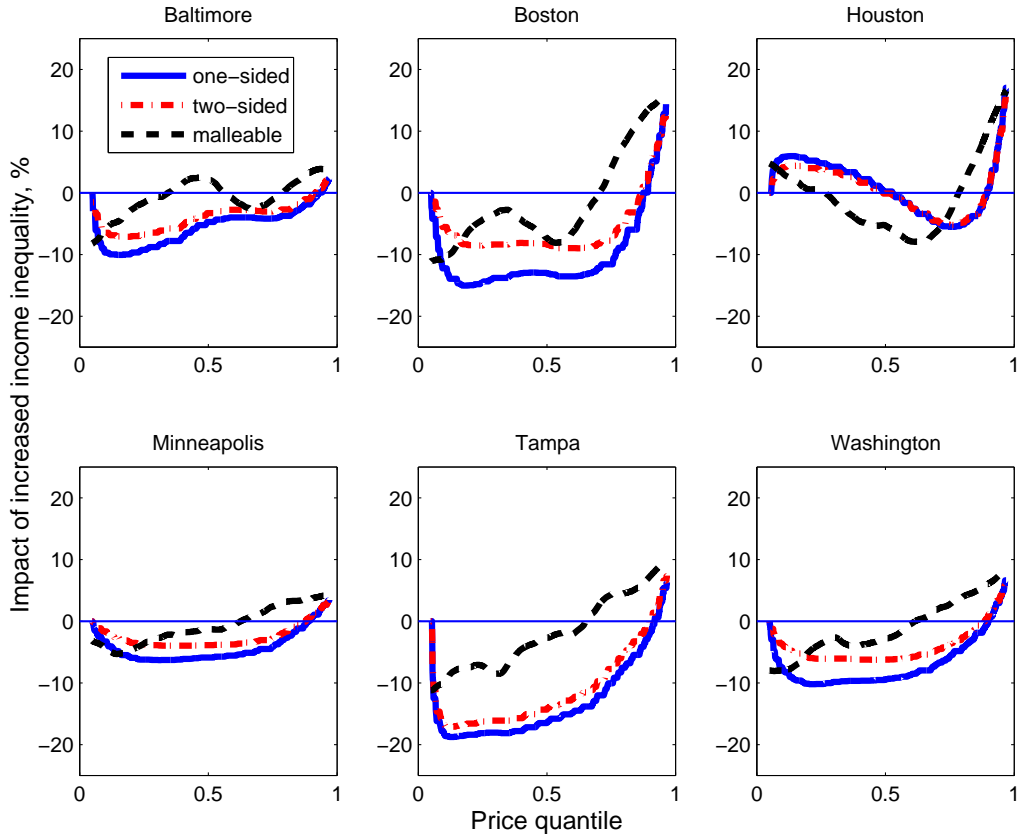


Figure 6: *Impact of increased income inequality on house prices/housing capital in alternative models.*

The equilibrium can be understood intuitively by considering the price gradient, which is, loosely, the price difference between “neighboring” houses in the quality distribution. The price gradient expresses how much households that inhabit a particular part of the quality distribution in equilibrium are willing to pay for the quality difference over next best house. This depends on their marginal rate of substitution between house quality and other goods, which in general depends on their level of wealth. The price level at any quantile in the distribution is the sum over all price gradients below.

The natural comparative static in our setup is order-preserving changes in income distribution, as these have an impact on prices without generating trading. The model yields a number of theoretical implications about the relation of income and house price distributions. (Thus it also yields results about the relation of the progressivity of income taxation and house price distribution). Any increase in income levels will increase both the level and dispersion of house prices. An increase in income inequality will decrease

house prices except (possibly) in a segment adjacent to the top. House prices at the top can go either way in response to an increase in income inequality, depending on the details of both supply and demand sides of the market.

As an illustration of the model, we obtained a theory-driven estimate for the impact of recent increases in income inequality on house prices in the US. The equilibrium conditions of the model allow us to estimate the housing quality distribution without imposing restrictions on the shapes of the distributions. Specifically, we asked how the 2007 price distribution would differ from the actual price distribution if income of every household would have grown at the actual mean rate since 1998. We found that the impact of increased inequality on prices has been modest but negative on average, and positive only at the top decile. This is because the cumulative impact of reductions in the price gradient at the bottom households, whose income growth did not keep up with the mean growth, dominates the positive effects almost all the way to the top of the distribution.

Our model opens the possibility for other applications. Most directly, it is a natural framework for studying how various housing and income subsidy schemes impact housing prices. The intuition of the price gradient makes it clear why housing subsidies targeted for the poor will not merely be capitalized into prices of low-quality housing but will spill upwards in the quality ladder. A serious empirical analysis of this issue will require the inclusion of non-owner-occupied housing. Other important and challenging directions include migration (choice between cities) and preference heterogeneity.

Both one-sided and two-sided matching, when assumed in “pure form” are abstractions and extreme ends of a spectrum of possible assumptions about the population of households in the local market. In a one-sided setup the local market is a closed economy, while in a two-sided setup all sellers are locals who move out of the economy, and all buyers are outsiders who move in. A model of a local housing market that would combine both type of trades—households who both buy and sell within the local market, and households that in and out for exogenous reasons (such as preference shocks), would of course be more realistic and more complicated.

The counterfactuals presented in this paper took the form of order-preserving changes in the exogenous income distribution. This was crucial for being able to apply the formulae derived under the assumption of no-trade equilibrium. The lack of trading also meant that price changes do not have welfare effects—they are merely changes in paper wealth. In order to have welfare effects, price changes have to generate trading. One interesting and challenging topic for further study is the welfare effects of regulations, for instance transaction taxes, which can be expected to distort the matching of houses and households. A particularly policy-relevant issue is the impact of repealing rent control, which would

result in a simultaneous supply and demand shock. Again, the impact on the entire distribution of housing prices is likely to be nontrivial. Our characterization of the (no-trade) equilibrium should be helpful even in applications that involve trading, although these will require more involved numerical methods.

## Appendix A: Proofs

### Proof of Lemma 1: Positive assortative matching by wealth and house quality

Define the reservation price  $\Delta(x'|x, y)$  by

$$u(x, y) = u(x', y - \Delta), \quad (22)$$

so that  $\Delta$  is the maximum price that a household with endowment  $(x, y) \in \mathbb{R}_+^2$  is willing to pay to switch to a house of alternative type,  $x'$ . Thus  $\{x', y + \Delta(x'|x, y)\}$  traces an indifference curve that goes through the endowment  $(x, y)$ ; it is strictly decreasing, and  $\Delta$  changes sign at  $x' = x$ . The reservation price of switching into a worse house is negative.

Consider two households,  $h = 1, 2$ , endowed with houses  $x_1 < x_2$ . There is trade between them if and only if their reservation prices for the trade sum up to something positive, i.e., if

$$\Delta(x_2|x_1, y_1) + \Delta(x_1|x_2, y_2) > 0. \quad (23)$$

We want to show that if  $x_1 < x_2$  and  $u(x_1, y_1) > u(x_2, y_2)$  then there has to be trade, which this rules out any violations of PAM in equilibrium. The wealthier household has a larger budget set, so PAM by wealth and house quality is equivalent with PAM by utility and house quality.

Consider a point  $\{x_2, y'_2\}$  where  $y'_2 = y_1 - \Delta(x_2|x_1, y_1)$ . This is along the indifference curve going through  $(x_1, y_1)$  and vertical in relation to  $(x_2, y_2)$ . Swapping positions between  $(x_2, y'_2)$  and  $(x_1, y_1)$  does not change utility so if  $y_2 = y'_2$  then  $\Delta(x_1|x_2, y'_2) + \Delta(x_2|x_1, y_1) = 0$ . If  $\Delta(x_1|x_2, y)$  is strictly decreasing in  $y$  then there will be trade if and only if  $y_2 < y'_2$ , which is equivalent to  $u(x_1, y_1) > u(x_2, y_2)$ . Differentiating (22) we obtain

$$\frac{d\Delta}{dy} = 1 - \frac{u_y(x_2, y)}{u_y(x_1, y - \Delta)}. \quad (24)$$

We know that  $u_y$  is positive and decreasing for a fixed  $x$ , but now the comparison is at two different levels of  $x$ , yet on the same indifference curve. Use  $\tilde{y}(x)$  to denote the indifference curve. Thus (24) is negative if

$$\frac{d}{dx} (u_y(x, \tilde{y}(x))) = u_{xy} + u_{yy} \left( \frac{d\tilde{y}(x)}{dx} \right) = u_{xy} - u_{yy} \left( \frac{u_x}{u_y} \right) \geq 0 \quad (25)$$



This is just the condition for the marginal rate of substitution  $MRS_{yx}$  to be decreasing in  $y$  (i.e., for  $MRS_{xy}$  to be increasing in  $y$ ), which holds for any quasi-concave utility function.

**Proof: Existence of equilibrium** Let's now discretize the house types, so that  $0 < x_0 < \dots < x_N < \infty$ . For brevity, we will refer to households endowed with a type- $k$  house as households of type  $k$ . Denote the mass of type- $k$  households with income equal or lower than  $\theta$  by  $F_k(\theta)$ . We assume that all of these conditional income distributions are continuous with full support  $[\underline{\theta}, \bar{\theta}]$ , where  $0 < \underline{\theta} < \bar{\theta} < \infty$ . The mass of type- $k$  households is  $m_k > 0$ , so  $F_k(\bar{\theta}) = m_k$ . Recall that we have normalized the total mass at  $\sum m_k = 1$ .

With discretized house types we require an additional assumption:

$$u(x_{k-1}, \bar{\theta}) > u(x_k, \underline{\theta}) \text{ for all } k = 1, \dots, N. \quad (26)$$

This means that distribution of autarky utility has overlap between neighboring household types. Together with Lemma 1, this implies that there must be trade between them, which in turn will guarantee that price increments are uniquely determined.

Lemma 1 shows that equilibrium utility is increasing in the quality of the house consumed. The full support of incomes and the overlap of autarky utilities imply that the consumption levels  $y$  of households that consume a house of type  $k$  must be distributed with full support on some interval  $[\underline{y}_k, \bar{y}^k] \in [\underline{\theta}, \bar{\theta}]$ . The unequivocally poorest and richest households do not have anyone to trade with, so  $\underline{y}_0 = \underline{\theta}$  and  $\bar{y}_N = \bar{\theta}$ .

The wealthiest household that consumes a type- $k$  house must be indifferent between trading up to the next house type. This indifference condition defines the price increments  $\Delta_k = p_k - p_{k-1}$  as functions of upper bounds  $\bar{y}_k$ ,

$$\gamma(\bar{y}_k | x_k, x_{k+1}) = \{\Delta_{k+1} : u(x_k, \bar{y}_k) = u(x_{k+1}, \bar{y}_k - \Delta_{k+1})\} \text{ for all } k = 0, \dots, N-1. \quad (27)$$

Note that  $\gamma$  is single-valued for all  $\bar{y}_k \in [\underline{\theta}, \bar{\theta}]$ , with image in  $(0, \bar{y}_k)$ .

The same indifference conditions can also be described in terms of the lower bounds of the consumption intervals, with  $u(x_k, \underline{y}_{k+1} + \Delta_{k+1}) = u(x_{k+1}, \underline{y}_{k+1})$ , so

$$\underline{y}_{k+1} = \bar{y}_k - \Delta_{k+1} = \bar{y}_k - \gamma(\bar{y}_k | x_k, x_{k+1}). \quad (28)$$

Thus the lower bounds as well as the price increments are uniquely determined by the upper bounds, so the equilibrium allocation can be described in terms of the  $\bar{y}_k$  alone, where  $k = 0, \dots, N-1$ .<sup>19</sup>

---

<sup>19</sup>Equivalently, the allocation can be described in terms of the lower bounds  $\underline{y}_k$  or prices  $\Delta_k$ , where  $k = 1, \dots, N$ . This choice of independent variables is a matter of convenience.

Demand for type- $k$  houses is the sum of demands from each household type. Consider type- $j$  households endowed with income  $\theta$ . They will consume a type- $k$  house if their wealth is in the same range as of those type- $k$  households who consume their endowment:  $p_j + \theta \in [p_k + \underline{y}_k, p_k + \bar{y}_k]$ . The bounding inequalities for these intervals can be written as

$$\begin{aligned}\theta &\leq \bar{y}_k + p_k - p_j \\ \theta &\geq \underline{y}_k + p_k - p_j = \bar{y}_{k-1} + p_{k-1} - p_j\end{aligned}\tag{29}$$

for  $k = 1, \dots, N$ . (Recall that  $\bar{y}_N = \bar{\theta}$ .) Total demand for type- $k$  houses is

$$\begin{aligned}Q_k(\bar{y}) &= \sum_{j=0}^N [F_j(\bar{y}_k + p_k - p_j) - F_j(\bar{y}_{k-1} + p_{k-1} - p_j)] \\ &= \sum_{j=0}^{k-1} [F_j(\bar{y}_k + (\Delta_{j+1} + \dots + \Delta_k)) - F_j(\bar{y}_{k-1} + (\Delta_{j+1} + \dots + \Delta_{k-1}))] \\ &\quad + F_k(\bar{y}_k) - F_j(\bar{y}_{k-1} - \Delta_k) \\ &\quad + \sum_{j=k+1}^N [F_j(\bar{y}_k - (\Delta_{k+1} + \dots + \Delta_j)) - F_j(\bar{y}_{k-1} - (\Delta_k + \dots + \Delta_j))],\end{aligned}\tag{30}$$

where  $\Delta_k$  are functions of  $\bar{y}$  as defined by (27) and (28). Define the excess demand functions as

$$Z_k(\bar{y}) = \max\{0, Q_k(\bar{y})\} - m_k,\tag{32}$$

where the infeasible negative demands allowed by (30) are removed. The nonnegativity constraint is binding if the upper bound  $\bar{y}_k$  is below the lower bound  $\underline{y}_k$  implied by  $\bar{y}_{k-1}$ .

Each household has zero net demand for housing units, because not trading amounts to demanding your own house. Thus  $Z_N(\bar{y}) = -\sum_{k=0}^{N-1} Z_k(\bar{y})$  and market equilibrium can be defined as

$$Z_k(\bar{y}) = 0 \text{ for } k = 0, \dots, N-1,\tag{33}$$

or, in vector form,  $Z(\bar{y}) = 0$ . This is a system of  $N$  equations in  $N$  unknowns.

Consider the best response of an imaginary "player" who attempts to minimize the absolute value of excess demand for type- $k$  houses with the sole instrument of choosing the upper bound  $\bar{y}_k$ . (The top house type, where  $\bar{y}_N = \bar{\theta}$  by construction, does not have a "player" representing it.) Use  $\bar{y}_{-k}$  to denote the vector  $\bar{y}$  without the  $k$ :th element. The best response function is

$$a_k(\bar{y}_{-k}) = \arg \min_{a \in [\underline{\theta}, \bar{\theta}]} |Z_k(a|\bar{y}_{-k})|.\tag{34}$$

It is straightforward to show that  $Z_k$  is continuous and increasing in  $a$ , but there may be a flat region. This happens when  $a$  is below the lower bound  $\underline{y}_k$  determined by  $\bar{y}_{-k}$ . Then demand for  $k$ -type houses is zero, and  $Z_k(a|\bar{y}_{-k}) = -m_k$ . However, the highest possible lower bound is strictly below  $\bar{\theta}$ , by (28) it is  $\bar{\theta} - \gamma(\bar{\theta}|x_{k-1}, x_k)$ , so  $Z_k$  must be strictly increasing in some interval  $[\theta, \bar{\theta}]$ . There are two possibilities

$$a_k(\bar{y}_{-k}) = \begin{cases} \bar{\theta}, & Z_k(\bar{\theta}|\bar{y}_{-k}) \leq 0 \\ \theta, & Z_k(\theta|\bar{y}_{-k}) = 0 \end{cases}, \quad (35)$$

where  $\theta \in (\underline{\theta}, \bar{\theta})$ . Thus  $a_k$  is single-valued but may not be able to eliminate excess supply for  $k$ . Hence  $a(\bar{y}) = (a_0(\bar{y}_{-0}), \dots, a_{N-1}(\bar{y}_{-(N-1)}))$  defines a continuous vector-valued function from  $[\underline{\theta}, \bar{\theta}]^N$  to itself. Thus, by Brouwer Fixed Point theorem, there exists a fixed point  $y^* = a(y^*)$ .

Next we prove, by contradiction, that a fixed point must be a market equilibrium. Suppose that  $y^* = a(y^*)$  but  $Z_k(y^*) < 0$  for some  $k$ , so  $y_k^* = \bar{\theta}$ . Now consider the highest such  $k$ , so that  $Z_h(y^*) = 0$  for all higher types  $h \in \{k+1, \dots, N-1\}$ . Since  $\underline{y}_{k+1} = \bar{\theta} - \gamma(\bar{\theta}|x_k, x_{k+1}) > \underline{\theta}$  (due to overlap of autarky utilities) at least some  $k+1$ -types are trading down, while no one among types  $k$  or lower is trading up to houses above  $k$ . Hence  $Z_h > 0$  for at least one  $h < k$ , a contradiction with (35). Intuitively, all demand for houses  $k+1$  and higher would have to come from households endowed with those houses. Since some of them are trading down, there must be excess demand for some  $h > k$ .

Thus any fixed point of  $Z$  indeed defines an equilibrium allocation of our model. Finally, since  $\gamma$  is strictly increasing, the equilibrium allocation is associated with unique price increments  $\Delta_k$ , so associated equilibrium prices  $p_k$  are unique up to an additive constant. (However, we have not shown that there is a unique equilibrium allocation).

## Appendix B: Data and preference inference

We use income and house price data from the American Housing Survey (AHS). House price is based on the survey question where respondents were asked to estimate the current market value of their house. We consider only homeowners, which amounts to assuming that rental housing forms an entirely separate market.

The choice of MAs was determined by the availability of data. The most populous MAs are covered only in national-level data (the National Surveys), where house prices are top coded at a common national threshold. This results in a disproportionate censoring of observations in the largest MAs. In those MAs where the national data is not decimated by top coding it suffers from small sample size, which unfortunately renders the national

surveys useless for our purposes. Mid-size cities are covered in separate "Metro surveys," which have been conducted at irregular intervals and which are less affected by the top-coding problem because they are done at city-level. Our sample includes all MAs from which there is metro survey data both in 2007 (the most recent year) and 1998 (the second most recent year for the MAs surveyed in 2007). Table B1 provides information about the number of observations, inequality, the rank correlation between income and house value, and population growth.

	#Obs		Gini		Gini		Gini		Rank corr.		Pop.
	1998	2007	perm. income	income	1998	2007	1998	2007	1998	2007	growth %
Baltimore	2306	1211	0.16	0.17	0.37	0.40	0.27	0.28	0.38	0.48	6.2
Boston	1960	1075	0.13	0.17	0.36	0.44	0.25	0.23	0.48	0.46	3.7
Houston	1963	1105	0.18	0.20	0.39	0.41	0.30	0.27	0.53	0.52	26.1
Minneapolis	2741	1524	0.14	0.16	0.33	0.36	0.23	0.23	0.50	0.49	10.5
Tampa	2106	1255	0.19	0.23	0.41	0.52	0.29	0.31	0.52	0.48	16.2
Washington	2356	1315	0.14	0.17	0.33	0.39	0.25	0.23	0.50	0.49	16.4

**Table B1.** *Number of observations, summary statistics for inequality (Gini coefficient) of permanent income, income, and house value, and for the Spearman rank correlation between income and house value, and population growth from 2000 to 2010 according to Census data. "Permanent income" is income kernel smoothed relative to house price.*

Figure 7 shows the relative prediction errors,  $\log(\hat{p}^{07}/p^{07})$ , by quantile of house value for the three MAs where the population has grown the least from 2000 to 2010. The solid lines show the prediction errors given the MA specific elasticity of substitution that minimizes the mean absolute percentage error (MAPE) in each MA. The elasticities are 0.42, 0.63, and 0.48 for Baltimore, Boston and Minneapolis, respectively. The resulting MAPEs are 4.6% (Baltimore), 2.8% (Boston), and 2.6% (Minneapolis). The model seems to provide a reasonable fit, especially in Boston and Minneapolis. The dashed lines show the prediction errors given our baseline assumption of an elasticity of substitution equal to 0.5.

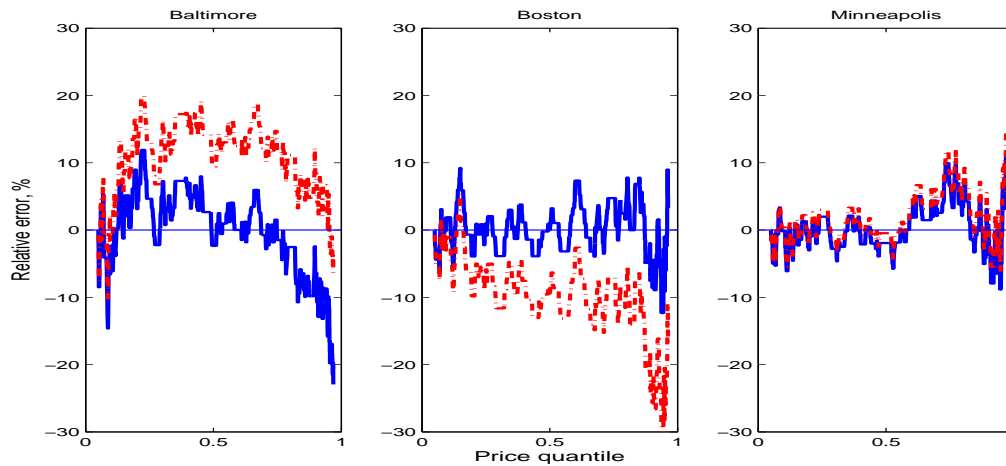


Figure 7: *Errors in predicted 2007 house prices.*

## References

- ALONSO, WILLIAM (1964). *Location and Land Use*. Harvard University Press, Cambridge, MA.
- BAJARI, PATRICK; CHAN, PHOEBE; KRUEGER, DIRK AND DANIEL MILLER (2010). A Dynamic Model of Housing Demand: Estimation and Policy Implications. *NBER Working Paper 15955*.
- BRAID, RALPH M (1981). “The Short-Run Comparative Statics of a Rental Housing Market.” *Journal of Urban Economics*, 10, pp. 286–310.
- CAPLIN, ANDREW AND JOHN LEAHY (2010). “A Graph Theoretic Approach to Markets for Indivisible Goods.” *NBER Working Paper 16284*.
- DAVIS, MORRIS AND JONATHAN HEATHCOTE (2007). “The Price and Quantity of Residential Land in the United States.” *Journal of Monetary Economics*, 54, pp. 2595–2620.
- DAVIS, MORRIS AND MICHAEL PALUMBO (2008). “The Price of Residential Land in Large US Cities.” *Journal of Urban Economics*, 63, pp. 352–384.
- DIAMOND, REBECCA (2012). “The Determinants and Welfare Implications of US Workers’ Diverging Location Choices by Skill:1980-2000.” *Harvard mimeo*.
- ECKHOUT, JAN; ROBERTO PINHEIRO, AND KURT SCHMIDHEINY (2013): “Spatial Sorting.” *Working paper*.
- EPPLER, DENNIS AND HOLGER SIEG (1999). “Estimating Equilibrium Models of Local Jurisdictions.” *Journal of Political Economy*, 107(4), pp. 645–681.

FUJITA, MASAHISA (1989). *Urban Economic Theory: Land Use and City Size*. Cambridge University Press, Cambridge, UK.

GALE, DAVID (1984). “Equilibrium in a Discrete Exchange Economy with Money.” *International Journal of Game Theory*, 13(1), pp. 61–64.

GLAZER, AMIHAI; VESA KANNIAINEN AND PANU POUTVAARA (2008). “Income Taxes, Property Values, and Migration.” *Journal of Public Economics*, 92, pp. 915–923.

GYOURKO, JOSEPH; CHRISTOPHER MAYER AND TODD SINAI (2006). “Superstar Cities.” *NBER Working Paper 12355*.

LANDVOIGT, TIM; MONIKA PIAZZESI AND MARTIN SCHNEIDER (2012). “The Housing Market(s) of San Diego.” *NBER Working Paper 17723*.

LEGROS, PATRICK AND ANDREW F NEWMAN (2007). “Beauty Is a Beast, Frog Is a Prince: Assortative Matching with Nontransferabilities.” *Econometrica*, 75(4), pp. 1073-1102.

LI, WENLI; HAIYOUNG LIU AND RUI YAO (2009). “Housing over Time and over the Life Cycle: A Structural Estimation.” *Federal Reserve Bank of Philadelphia Working Paper*.

MÄÄTTÄNEN, NIKU AND MARKO TERVIÖ (2009). “Income Distribution and Housing Prices: An Assignment Model Approach.” *HSE Working Paper*.

MORETTI, ENRICO (2013). “Real Wage Inequality.” *American Economic Journal: Applied Economics*, 5(1), pp. 65-103.

ORTALO-MAGNÉ, FRANÇOIS AND SVEN RADY (2006). “Housing Market Dynamics: On the Contribution of Income Shocks and Credit Constraints.” *Review of Economic Studies*, 73, pp. 459-85.

ORTALO-MAGNÉ, FRANÇOIS AND ANDREA PRAT (2010). “Spatial Asset Pricing: A First Step.” *Working paper*.

QUINZII, MARTINE (1984). “Core and Competitive Equilibria with Indivisibilities.” *International Journal of Game Theory*, 13(1), pp. 41–60.

RÍOS-RULL, JOSÉ-VICTOR AND VIRGINIA SANCHEZ-MARCOS (2008): “An Aggregate Economy with Different House Sizes.” *Journal of The European Economic Association*, P&P, 6(2-3), pp. 705-14.

SATTINGER, MICHAEL (1979). “Differential Rents and the Distribution of Earnings.” *Oxford Economic Papers*, 31, pp. 60-71.

SATTINGER, MICHAEL (1993). “Assignment Models of the Distribution of Earnings.” *Journal of Economic Literature*, 31, pp. 831-880.

SHAPLEY, LLOYD AND HERBERT SCARF (1974). “On Cores and Indivisibility.” *Journal of Mathematical Economics*, 1(1), pp. 23–27.

TERVIÖ, MARKO (2008). “The Difference that CEOs Make: An Assignment Model Approach.” *American Economic Review* 98(3), pp. 642–668.

TEULINGS, COEN N (1995). “The Wage Distribution in a Model of the Assignment of Skills to Jobs.” *Journal of Political Economy*, 102, pp. 280–315.“

TIEBOUT, CHARLES M (1956). A Pure Theory of Local Expenditure.” *Journal of Political Economy*, 64, pp. 416–424.

VAN NIEUWERBURGH, STIJN AND PIERRE-OLIVIER WEILL (2010). “Why Has House Price Dispersion Gone Up?” *Review of Economic Studies*, 77(4), pp.1567–1606.