Income Distribution and Housing Prices: An Assignment Model Approach*

Niku Määttänen
ETLA and HECER†

Marko Terviö
Aalto University and HECER‡

February 9, 2010

Abstract

We present a framework for studying the relation between the distribution of income and the distribution of housing prices that is based on an assignment model of households with heterogeneous incomes and houses of heterogeneous quality. The equilibrium distribution of prices depends on both distributions in a tractable but nontrivial manner. We infer the unobserved distribution of quality from the joint distribution of income and housing prices, and use it to generate counterfactual price distributions under counterfactual income distributions. We apply the model to estimate the impact of recently increased income inequality on the distribution of house prices in the Helsinki Metropolitan region. We find that recent increase in income inequality caused house prices to be lower than if income growth had been uniform across the population. JEL: D31, R21.

*We thank Pauli Murto, Otto Toivanen, and Juuso Välimäki for useful suggestions.

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

‡Aalto University School of Economics and Helsinki Center for Economic Research. *http://hse-econ.fi/tervio/*

1 Introduction

One question raised by the recent increases in income inequality is whether it has had an impact on the distribution of housing prices. Could it be that ever higher top incomes just lead the rich to bid the price of best locations ever higher? To what extent get higher top incomes capitalized in house prices? It has been argued that the increase in consumption inequality has been less than the increase in income inequality for essentially this reason (Moretti 2009). We present an assignment model framework to study this question and, in the process, also generate theoretical insights on the relation of income inequality and house prices.

In our view, from a distributional perspective, a central feature of the housing market is that housing is not a fungible commodity but comes embedded in indivisible and heterogeneous units. The supply of houses is more or less fixed in short and medium run, and in terms of the most desirable locations, also in the long run. Another key feature is that housing is a normal good that takes up a large part of household expenditure. Thus, unlike in standard assignment models, we do not assume transferable utility. Our model can be summarized as an assignment model with income effects. The market consists of a population of households who each own one house and each wish to live in one house. Even though there are no complementarities in the usual sense, this setup results in positive assortative matching: wealthier households end up living in the higher-quality houses. To focus on the impact of changes in the income distribution we assume that households have the same preferences: here the only reason why the wealthy live in better houses is that they can better afford them. The equilibrium distribution of house prices depends on the shapes of both the distribution of house quality and the distribution of income in a tractable although nontrivial manner. With this model, we are equipped to study questions such as the impact of changes in income distribution on the distribution of housing prices.

In the general version of our model, the joint distribution of houses and income (non-housing wealth) is arbitrary, which results potentially in a lot of trading between households. Equilibrium prices depend on the joint distribution of endowments, not just on the (marginal) distributions of income and house quality. For our application, we interpret the observed house prices as the equilibrium prices that would emerge after all trading opportunities have been exploited.

Under this assumption we ask what distribution of unobserved house quality, together with the observed distribution of incomes, would give rise to the observed price distribution as the equilibrium outcome of our model. We find that, under a suitably parametrized CES utility function, this unobserved distribution is stable over time, which fits with our intuition that the quality distribution is fixed in the short run. We then use the inferred distribution of house qualities and the preferred utility parametrization to generate counterfactuals to answer the motivating question of the paper.

In our empirical application we use data from the Helsinki metropolitan region in 1998 and 2004, where there was a significant increase in income inequality. We consider a counterfactual income distribution for 2004 where all incomes grow uniformly since 1998 at the same rate as the actual mean income. (I.e. the shape of the counterfactual distribution is the same as the actual shape in 1998). This counterfactual generates house prices that are on average about 2% higher, implying that the increase in inequality has resulted in lower house prices than would have prevailed under uniform income growth. Only at the top decile has increased income inequality had a positive impact in higher prices.

The reason why uniform income growth would have lead to higher prices at the bottom of the quality distribution is intuitive: if low-income households had higher incomes they would use some of it to bid for low-quality houses. However, in a matching market with positive sorting, 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 price difference between two "neighboring" houses in the quality distribution is pinned down by how much the households at the respective part of the income distribution are willing to pay for the quality difference. Any price changes at the bottom of the distribution spill upwards, and this downwards pressure on prices from the bottom of the distribution counteracts the local increase in willingness-to-pay among better-off households whose incomes are now higher than they'd been under uniform growth. In principle, it would therefore be possible for all house prices to go down in response to an increase in inequality. In our data we find that, while incomes have grown more than average in the top half of the distribution, the net impact of uneven income growth on house prices has been positive only at the top $5-10\%$ of the distribution.

In our setup, households own the houses to begin with, so they are also the beneficiaries and losers of any changes in house prices. For the empirical application we assume that only the shape of the distribution has changed over time, but that the ranking of households by income has stayed the same. In this case changes in house prices are not changes in wealth in any reasonably defined sense.[1] More generally, if a household's position in the income distribution changes then it would want to change houses, and would on net benefit or lose based on how others' incomes have changed. For example, a household whose income and rank in income distribution both increase will lose some of the benefit if others' incomes have risen too. It will have to give up some of the increase in income in order to move up in the ranking of house quality. Trading will then involve a net transfer of non-housing wealth to someone who is now lower in the income distribution. The distributional impact of a change in house prices is that the newly rich lose out and the no-longer quite so rich gain, relative to a "naive" measure where housing is assumed to be a fungible good with a single unit price.

In the next section we discuss related literature. In Section 3 we present the general version of the model. In Section 4 we show how the model can be used for inference under the assumption that observed prices reflect the post-trade or "steady state" allocation of our model. Section 5 presents the empirical application, and Section 6 concludes.

2 Literature

Our model is an otherwise standard assignment model but with non-transferable utility. Assignment models are models of matching markets with symmetric information and no frictions; for a review see Sattinger (1993). Perfect competition is achieved by assuming a continuum of types on both sides of the market. There is no room for "bargaining" as all agents have a continuum of arbitrarily close competitors. Assignment models typically include an assumption of a complementarity in production, which results in assortative matching and equilibrium prices that 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

[1] See also "Housing Wealth Isn't Wealth" by Buiter 2008. In our setup, only increases in house quality could result in aggregate increases in (reasonably defined) housing wealth.

the productive complementarity is between job types and worker types, as in Sattinger (1979) and Teulings (1995), or between workers themselves in a team production setting, as in Kremer (1993). In our application there is no complementarity, but equilibrium nevertheless involves assortative matching by income and house quality, due to housing being a normal good. We don't restrict the shapes of the distributions, and our nonparametric method for inferring the unobserved type distribution is similar to Terviö (2008).

Matching models have long been applied to the housing market from a more theoretical perspective, although it is perhaps more accurate to say that housing has often been used in theoretical matching literature 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 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. Miyagawa (2001) extends the model by adding a second, continuous good, i.e., "money." He shows that the core assignment of houses can be implemented with a set of fixed prices for the houses. In Miyagawa's model utility is quasilinear, so there is no potential for income effects. In our setup there are both indivisible and continuous goods, and utility is concave in the continuous good.

There is a long tradition in explaining heterogeneous land prices in urban economics, going back to Von Thünen (1826) and Alonso (1964). In urban models the exogenous heterogeneity of land is due to distance from the center. The focus in urban economics is on explaining how land use is determined in equilibrium, including phenomena such as parcel size and population density. In modern urban economics, see Fujita (1989), there are also some models with income effects. Heterogeneity of land is modeled as a transport cost, which is a function of distance from center, and price differences between locations are practically pinned down by the transport cost function.

Much attention has also been devoted to the question of endogenous public good provision, in the tradition of Tiebout (1956). Epple and Sieg (1999) estimate preference parameters in a structural model where the equilibrium looks like assortative matching by income and public good quality, although the latter is a choice variable at the level of the community. Glazer, Kanniainen, 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 better land, allowing the poor to consume better land than before.

Most of the dynamic macroeconomic models with housing assume that housing is a homogenous malleable good. In any given period, there is then just one unit price for housing. An exception is the property ladder structure that is used in Ortalo-Magne and Rady (2006) and Sanchez-Marcos and Rios-Rull (2008), where there are two types of houses: relatively small “flats” and bigger “houses”. For our purposes, such a distribution would be far too coarse. In general, the macro 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. By contrast, we focus on the cross-sectional and distributional aspects of the housing market, and abstract away from the time-series aspect.

Van Nieuwerburgh and Weill (2009) study house price dispersion across US cities using a dynamic model, where there is matching by individual ability and regional productivity. Within each location housing is produced with a linear technology, but there is a resource constraint for the construction of new houses. This causes housing to become relatively more expensive in regions that experience increases in relative productivity. Houses are non-tradable across cities while labor is mobile, so intuitively the result is similar to the Balassa-Samuelson effect. In their calibration Van Nieuwerburgh and Weill find that, by assuming a particular increase in the dispersion of ability, they can reasonably well generate the observed increase in wage dispersion and the (larger) increase in house price dispersion across cities. 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 result in more bidding for the scarce land, thus leading the price of houses in the superstar city to go up.

One step in our empirical application is that we estimate the elasticity parameter of a constant elasticity of substitution utility function for housing and other consumption. Therefore, our paper is also related to studies that estimate the intratemporal elasticity of substitution between housing and other consumption. A recent example of a paper that uses a structural approach for that is Li, Liu and Yau (2009). They estimate the preference parameters by fitting a life cycle model with housing to both aggregate time series and cross-section US data. However, as far as we know, we are the first to exploit changes in the distribution of housing prices to estimate household preference parameters. This is possible precisely because in our model housing prices are in general a non-linear function of housing quality.

3 Model

We consider a one-period economy where a unit mass of households consume two goods, housing and a composite good. Preferences are described by a standard concave utility function u where both goods are normal. 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 consumes exactly one house. Households have the same preferences but different endowments. Household income, denoted by θ, is interpreted as the endowment of the composite good.

There are no informational imperfections in the model, or other frictions besides the indivisibility of houses. There is a continuum of households and houses, so there is also no market power as every house has "infinitely many" arbitrarily close substitutes (each with a different owner).

Figure 1 depicts this economy. A household endowment (x, θ) is described by a point in $[0, 1] \times \mathbb{R}_+$, where the horizontal dimension represents $F_x(x)$, the quantile in the distribution of house quality, and the vertical dimension represents the amount of composite good. As preferences are homogeneous, the same indifference map applies to all households.

We use the composite good y as the numeraire, while p is the equilibrium price function for house quality. Budget constraints are downward sloping curves, because house prices must be increasing in quality. Figure 1 depicts the budget curve of a household endowed with income θ and a house of quality $\tilde{x}$, it is defined by $\theta + p(\tilde{x}) = y + p(x)$, where y and x are the consumed values. Note that household wealth—the left side of the budget constraint—is endogenous, because the value of the endowment depends on p.

We assume that the lowest-quality (occupied) unit of housing, of quality $x_0 > 0$, is also available as an outside option at an exogenous price $p_0 \in (0, \theta_0)$, where θ_0 is the income of the poorest household. The joint distribution of endowments (x, θ) is assumed continuous and without gaps in its support, which implies that the quantile functions of θ and x are continuous and strictly increasing. (The quantile function at $x(i)$ is x such that $i = F_x(x)$). All households with an endowment on the same budget curve trade to the point where the budget curve is tangent to an indifference curve.

Among all households endowed with a house of a given quality $x(i)$ there exists a unique income level $\theta^*(i)$ at which the household does not trade. These no-trade endowments have a particularly helpful role in the model. The increasing curve in Figure 2 depicts the no-trade endowments under equilibrium prices. The starting point of the no-trade curve is necessarily $\{x_0, \theta_0\}$, the endowment of the unambiguously poorest households in this economy, who have nothing to offer in exchange. Households below the no-trade curve are endowed with a relatively high quality house and trade down in order to increase their consumption of the composite good, vice versa for households above the curve.

The no-trade curve is continuous, but not necessarily monotonic. To see this, consider a household with an arbitrary endowment $\{x, \theta\}$. If there existed only one other household, what should its endowment be for these two households to engage in trade? If both goods were continuously divisible then the answer would be whenever the marginal rates of substitution differ between the two households—which, in a generic sense, is always the case. However, when one of the goods is indivisible this is not enough. Even if the household that is endowed with the lower quality house is endowed with more money, it may not be possible to trade. Intuitively, as the households must swap their houses pairwise, the owner of the lower quality house must make a net payment in composite good (money). But if she already has lower utility than the owner of the high quality house, she may not have enough money for both households to be better off after swapping houses; see Figure 3 for the illustration. The point is that each endowment comes with a trading set, which includes all those endowments with which mutually profitable trade is possible. In equilibrium, it must be the case that no point on the no-trade curve is included in the trading set of another point on the curve. Together with the full support assumption this implies that the no-trade curve is continuous, with equilibrium utility (but not necessarily consumption y) increasing along it. As we show in the Appendix, trade between two households under CES utility is possible if and only if the one endowed with the better house has lower autarky utility.

Equilibrium consists of a price function p for houses and a matching of households to houses. The assumption that housing is a normal good results in assortative matching: of two households, the wealthier will live in the better and more expensive house. The twist here is that wealth is endogenous, because the value of the endowment depends on p. Hence equilibrium matching is not obvious as it depends both on the preferences and on the joint distribution of $\{x, \theta\}$. The resource constraints require that, for every i, the mass of households with an endowment below the budget curve that contains $\{x(i), \theta^*(i)\}$ is equal to i, the mass of houses of quality $x(i)$ or less.[2] In general, this leads to a rather complicated partial differential equation. By discretizing the house types, the equilibrium can still be solved numerically using standard recursive methods. However, in our empirical application we bypass this problem by assuming that the observed allocation is the equilibrium (post-trade) allocation.

Absentee landlord model

Before considering our main case, the post-trade economy, it is instructive to consider the case with "absentee landlords," i.e., an otherwise similar economy but where all houses are initially owned by competitive outside sellers.[3] Denote the distribution of household wealth by $F_w(w)$, and its inverse by $w(i)$. (The reason for treating wealth as distinct from income will become clear in the next section.) Consider the household at quantile i of the wealth distribution. From the fact that equilibrium must involve positive assortative matching by wealth and house quality we know that p must result in every household buying a house of the same quality rank as is their rank in the wealth distribution, so that

$$i = \arg\max_{j \in [0,1]} u\left(x\left(j\right), w\left(i\right) - p\left(j\right)\right) \tag{1}$$

[2]Note that, since the quality of houses is fixed, the distribution of x is the same for endowments and consumption. By contrast, for consumption of the composite good, y, only its mean value must match that of the endowment, θ.

[3]This is the standard assumption in urban economics, see e.g. Fujita (1989).

must hold for all $i \in [0,1]$. The associated first-order condition, $u_x x' - u_y p' = 0$, defines an ordinary differential equation for the equilibrium price:

$$p'(i) = \frac{u_x (x(i), w(i) - p(i))}{u_y (x(i), w(i) - p(i))} x'(i). \tag{2}$$

This price gradient is the key equation of the 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 sellers' 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 housing market). The price level at quantile i is the sum of the outside price $p(0)$ and the integral over all price gradients below i.

The intuition behind (2) is that the price difference between any neighboring houses in the quality order depends on how much the relevant households—i.e., those located at the respective part of the income order—are willing to pay for the respective quality difference. This depends on their marginal rate of substitution between house quality and other goods, which in turn may depend on their level of wealth. Note that the equilibrium price at any quantile i depends on the distribution of housing quality and income at all quantiles below i. Hence changes at any part of the distributions spill upwards in the price distribution, but not vice versa.

The asymmetry in the direction of price spillovers can be understood by considering the problem between "neighboring" households. The binding outside opportunity for any (inframarginal) household in equilibrium is that they must not want to trade to the next best house. Whatever happens above in the distribution of income (or house quality) has no impact, as long as the household itself remains in the same quantile.

It is worth noting that, in the light of our model, the claim that an increase in income inequality must lead to an increase in the prices of best houses (or land rents of most desirable locations) is incorrect. To see this, suppose some wealth is redistributed from poor to rich. It is true that this will increase the local "price gradient" (2) at the top quantiles, as the willingness-to-pay for extra quality goes up for the rich. But, for the same reason, the local price gradient at bottom quantiles will then go down. Due to the upwards spillover in price changes, it is in fact possible for all house prices to go down in response to an increase in inequality.

Post-trade model

Suppose the economy was initially at some arbitrary continuously distributed endowment, and then had the chance to engage in trade. As we saw above, the equilibrium allocations should be located on a no-trade curve, where there is perfect rank correlation between house quality and total wealth. Now, after trading but before consumption takes place, the prevailing equilibrium prices can be interpreted as the no-trade prices that enforce the equilibrium allocation. Mathematically this means that matching becomes very simple—it is positively assortative by observed wealth $p(x) + y$ and house quality x. This interpretation will be useful when we observe the distributions of p and y and wish to infer the unobserved x. In the empirical application we assume that the observed prices correspond to the equilibrium prices that emerge after all trading-opportunities have been exhausted. We think this is a reasonable interpretation of data because only a small fraction of households trade houses in a given period. In terms of the model, this amounts to using $\theta(i) + p(i)$ to replace $w(i)$ in (1). In other words, the no-trade prices are equivalent to the equilibrium prices that would result if houses were sold by absentee landlords and the wealth of each household i would happen to be $w(i) = \theta(i) + p(i)$.

Steady state interpretation

Suppose the population consists of household dynasties where each generation lives for one period, bequests its housing for the next generation, and has constant income, θ. Houses are durable and must be owned by the occupant. A generation only cares about its own utility, but the generations are linked by the houses, which are left for the next generation of the dynasty, and by the income level θ, which is a fixed characteristic of the dynasty. In steady state there can be no trading opportunities, so it must involve positive assortative matching by income and house quality. In steady state the role of house prices is then merely to enforce the no-trade equilibrium, so that, again, $w(i) = \theta(i) + p(i)$ in (1) and (2).

The case with CES

For the empirical application we assume CES utility,

$$u(x, y) = (\alpha x^\rho + (1 - \alpha) y^\rho)^{\frac{1}{\rho}}, \quad \text{where } \rho < 1 \text{ and } \alpha \in (0, 1), \tag{3}$$

with the Cobb-Douglas utility defined in the usual fashion at $\rho = 0$. The equilibrium price condition (2) becomes

$$p'(i) = \frac{\alpha}{1-\alpha}\left(\frac{w(i) - p(i)}{x(i)}\right)^{1-\rho} x'(i). \tag{4}$$

With the post-trade interpretation, $w(i) - p(i) = \theta(i)$, this can be solved as

$$p(i) = p_0 + \frac{\alpha}{1-\alpha}\int_0^i \left(\frac{\theta(s)}{x(s)}\right)^{1-\rho} x'(s)\,\mathrm{d}s. \tag{5}$$

Example: Pareto distributions Equilibrium prices have a closed-form solution in our model only under specific assumptions. Our empirical applications do not require such a solution, but they allow for tractable theoretical extensions. Here we present an example of a closed-form solution.

Assume that both income and house quality follow Pareto distributions, so that θ˜Pareto(θ_0, η) and x˜Pareto(x_0, γ), where $w_0 > p_0$ and $\eta, \gamma > 2$. The associated quantile functions are $\theta(i) = \theta_0 (1-i)^{-\frac{1}{\eta}}$ and $x(i) = x_0 (1-i)^{-\frac{1}{\gamma}}$. If utility takes the CES form (3), then under the post-trade interpretation, the equilibrium prices (4) can be solved in closed form:

$$p(i) = p_0 + \frac{\alpha}{1-\alpha}\frac{\xi}{\gamma}\theta_0^{1-\rho}x_0^{\rho}\left((1-i)^{-\frac{1}{\xi}} - 1\right) \tag{6}$$

where $\xi \equiv \frac{\eta}{1-\rho(1-\eta/\gamma)}$. This means that prices are distributed according to a Generalized Pareto Distribution. The expenditure share of housing $a(i) = p(i)/(p(i) + \theta(i))$ can be shown to have a limiting value at 1 if $\rho > 0$ and at 0 if $\rho < 0$. In the knife-edge Cobb-Douglas case it is

$$a(1) = \frac{\alpha\eta}{\alpha\eta + (1-\alpha)\gamma}. \tag{7}$$

The expenditure share a is then everywhere increasing if $a(1) > a_0 \equiv p_0/(p_0 + \theta_0)$ (decreasing if $a(1) < a_0$). If housing at the extensive margin can be created at constant marginal cost then the poorest household faces in effect linear prices and it is reasonable to assume that $a(0) = \alpha$. In this case the expenditure share of housing is strictly increasing in income if and only if $\eta > \gamma$, i.e., when the variance of wealth is lower than the variance of house quality.

This example illustrates how one cannot expect the expenditure shares to be constant across income levels, even if utility function takes the Cobb-Douglas form. The expenditure share of housing is not directly given by preferences because the prices faced by the consumers are nonlinear. The standard CES result that expenditure shares are independent of income is based on all goods being fungible, so that there is essentially just one type of housing, and households consume different amounts of it.

Example: Degenerate wealth distribution Suppose all households have the same wealth level $\bar{w}$ and preferences are Cobb-Douglas. Now prices must make every household indifferent between every housing unit. Then (4) simplifies to

$$p(i) = \bar{w} - (\bar{w} - p_0) \left(\frac{x_0}{x(i)} \right)^{\frac{\alpha}{1-\alpha}} \tag{8}$$

where $\bar{w} > p_0$ must be assumed. This (admittedly contrived) example provides the simplest demonstration for why the expenditure share of housing cannot be expected to be constant even if preferences are the same for all – some households simply must end up with the lower quality houses and for this they must be compensated with higher consumption of other goods.

4 Inference

The model can be used to infer the unobserved distribution of x from the observed relation between household income and house prices, if one knows or is willing to assume a utility function. Given u, $p(i)$, and $w(i)$, we can infer the distribution of x up to a constant. This is done by treating x as the unknown in the differential equation (2), while normalizing the constant $x(0) = x_0$, e.g. $x_0 = 1$. A number of interesting counterfactuals, where the normalization of x_0 is without loss of generality, can be generated with this inferred distribution.

The continuous model has the desirable feature that equilibrium prices are unique. However, it is useful to also understand the equilibrium in a discrete model, as we use the discrete formulae in our data application. (The continuous model can also be obtained as the limiting case of the discrete model with a large number of households and houses.) In the discrete model with exogenous wealth the equilibrium conditions for prices are

$$u(x_h, w_h - p_h) \geq u(x_{h'}, w_h - p_{h'}) \quad \text{for all } h, h'. \tag{9}$$

Thanks to assortative matching, only the constraints of the form $h' = h - 1$ are binding. If u has a closed-form inverse with respect to its second argument then p_h can be expressed as a function of $w_h, p_{h-1}, x_h, x_{h-1}$. Denoting this inverse as Y, where $Y(u(x,y),x) = y$, the price formula is

$$p_h = w_h - Y\left(u\left(x_{h'}, w_h - p_{h'}\right), x_h\right). \tag{10}$$

Discreteness results in match-specific rents, so here, for simplicity, we have assumed that sellers get all of them. (In our data, this makes no practical difference to the results compared to assuming that buyers get all match-specific rents.)

With CES utility the discrete price formula becomes

$$p_h = w_h - \left((w_h - p_{h-1})^\rho - \frac{\alpha}{1-\alpha}\left(x_h^\rho - x_{h-1}^\rho\right)\right)^{\frac{1}{\rho}}. \tag{11}$$

Denoting $\tilde{x} = x^\rho$ the price formula can be inverted and solved for

$$\tilde{x}_h = \tilde{x}_0 + \frac{1-\alpha}{\alpha}\sum_{n=1}^{h}\left[(w_n - p_{n-1})^\rho - (w_n - p_n)^\rho\right] \tag{12}$$

which includes an undefined constant of integration $\tilde{x}_0$. Note that when we infer x then the value of $\alpha \in (0,1)$ is without consequence for all potentially observable variables. This is because, in the utility function, changing α is equivalent to changing the units of x. Since x is not observed but only inferred within the model, this merely changes the estimated units of x, but has no impact on the monetary counterfactuals that we are interested in.

In the end, the purpose of inferring the distribution of an abstract quality unit is that it allows us to construct interesting counterfactuals. Suppose we have inferred the distribution of x based on the actual data on θ and p. Using a counterfactual income distribution $\tilde{\theta}$ we can then generate the counterfactual distribution of house prices, by combining $\tilde{\theta}$ and x in the (discrete equivalent of) equilibrium price relation (5). Note however, that as p_0 is exogenous to the model, our model only explains the differences in prices relative to the marginal unit of housing $p - p_0$. In the counterfactuals that follow the lowest price is always taken to be the lowest value in the data.

5 Empirical application

We now illustrate how the model can be used to infer household preference parameters based on changes in the distributions of income and housing prices. We use income and housing price data from the Metropolitan area of Helsinki. The inference is helped by the fact that Finnish income distribution has recently changed substantially. Finland had the largest increase in income inequality from mid-1990s to mid-2000s of all OECD countries.[4]

5.1 Data and smoothing

Our data is the two most recent Wealth Surveys by Statistics Finland, from 1998 and 2004. The Survey has detailed portfolio information from about 2500 households. We include both homeowners and market renters in the Metropolitan area of Helsinki, but we exclude households in public rental housing.[5] That leaves a sample of 483 households in the 1998 survey and 583 households in 2004.

We use just two variables. The first is total disposable income during the last year. This measure includes transfers and taxes. As above, we denote it by θ. The second variable we need is the current market value of the house, p. In the survey, respondents were asked to estimate the current market value of their house. To make the units of measurement comparable with the price of housing we convert income and rents to present value form by dividing by the discount rate. For brevity, we refer to the present value of housing expenditure as "house price".

For both 1998 and 2004, we observe the joint distribution of income θ and house price p. As we now use the post-trade version of the model, we need to estimate a relationship that reduces θ and p to a common increasing order. The two variables are not perfectly rank correlated in the data. (The rank correlation is 0.57 in 1998 and 0.58 in 2004.) In order to smooth the data we non-parametrically estimate $\theta(i)$ as $E[\theta|F_p(\theta) = i]$. The rationale for smoothing incomes instead of house prices is that yearly income is likely to be a noisy proxy for permanent income, while the current house value is presumably a relatively stable measure. We denote the smoothed distributions by $\{\theta^t, p^t\}$, $t \in \{98, 04\}$. Figure 4 displays the joint distributions together with the kernel regressed relationships $E[\theta|F_p(p) = i]$, which is our preferred smoothing method. Clearly, there is a strong positive relationship between the rank of the house value

[4] See Figure 1.2 in "Growing Unequal? Income Distribution and Poverty in OECD Countries", OECD, 2008.

[5] Most of public rental dwellings are part of social housing where the tenants are selected on the basis of social and financial needs, and the rents are regulated. There have been no major changes in the supply or means-testing criteria of public housing between 1998 and 2004.

and household's income. We also report the results using the converse smoothing, where house value is estimated as a function of income, see Figure 5. If income is more noisily measured, as we believe, then this converse smoothing is prone to be "too smooth."

The left panel of Figure 6 displays the distributions of income relative to its mean in 1998 and 2004 on log scale. There is a significant increase in income inequality from 1998 to 2004. The Gini coefficient for household incomes in our sample increases from 0.37 to 0.40. For our estimated "permanent income" the corresponding increase is from 0.21 to 0.23. It is this change in the income distribution, that allows us to infer preferences. The right panel of Figure 6 displays the equivalent picture for house prices. In relative terms, expensive houses have become more expensive and cheap houses cheaper.

5.2 Inferring preferences

We consider the CES-utility function specified in (3) and try to infer the elasticity parameter ρ. Given ρ and $\{\theta^{98}, p^{98}\}$, we first infer the quality distribution in 1998, denote it by x^{98}, using the inference formula in (12). In doing so, we set $w(i)$ in the formula at $w^{98}(i) = \theta^{98}(i)/r + p^{98}(i)$. Notice that the the fact that θ^{98} and p^{98} are perfectly rank correlated means that w^{98} and p^{98} are perfectly rank correlated as well. The interest rate r is needed to make the annual monetary income comparable to the house price. We fix the interest rate at $r = 0.032$, which was the average private rental return in Helsinki in 2007.[6] Changing r over a reasonable range ($2-5\%$) does not substantially change our inferred elasticity parameter ρ or the results from the counterfactual experiments that we present below.[7]

We then use the pricing formula (11) to predict the 2004 housing price distribution given w^{04} and x^{98}. In other words, assuming that the quality distribution x is fixed, we ask what would be the predicted price distribution in 2004 given the 2004 income distribution. Recall that the model does not explain the absolute level of housing prices but rather the difference

[6] Ilkka Lehtinen, Statistics Finland, unpublished seminar presentation.

[7] The inferred elasticity parameter does not change much even if we assume that the interest rate is different in 1998 and in 2004. This is because we try to match the entire distribution of housing prices, not just the average level. By contrast, measuring the proper interest rate is crucial in empirical work that tries to estimate housing demand elasticities by using time variation in some aggregate house price index.

to the price of the lowest quality house. In the counterfactuals we set the price of the lowest quality house equal to the actual lowest value. We assume that market prices would reflect the post-trade equilibrium, and that the ranking of households by income would not be affected, so that each household would still prefer to consume its endowment. This prevents the actual p from appearing in the counterfactual, which would lead to a circularity.

We compare the model's predicted housing price distribution to the empirical 2004 distribution. We repeat this exercise of inferring the quality distribution from 1998 data and comparing predicted and empirical housing price distributions for different values of ρ. Our preferred elasticity parameter is the one where the mismatch between the empirical and predicted 2004 housing price distribution is the smallest.

Figure 7 displays the empirical 2004 price distribution and the predicted price distributions for different values of ρ. Mean absolute error is minimized by selecting $\rho = 0.20$.[8] Our inferred value for the elasticity parameter is in line with studies that use aggregate time series data, but is substantially higher than in many studies using household data (see e.g. Li et al., 2009, and the references therein). As always in structural estimations, the interpretation of the parameter depends on the specifics of the model. The fixed distribution of indivisible houses is a friction that has not been taken into account before, so our estimate of ρ is not directly comparable with those obtained in previous studies.

5.3 Empirical Results

The key question we set out to ask was how changes in income distribution influence housing prices. We now apply our methodology to answer this question for our specific data set. First we compute the effect that the increased income inequality from 1998 to 2004 has had on housing prices in Helsinki metropolitan area. Specifically, given $\rho = 0.20$ and the inferred quality distribution, x^{98}, we compute the equilibrium housing price distribution that would be associated with an income distribution that is obtained by changing the 1998 income distribution proportionally to match the mean of 2004 incomes. We then compare this counterfactual price distribution with the price distribution that is obtained by plugging in the empirical distribution,

[8] Under converse smoothing the best fit is provided by $\rho = 0.08$.

θ^{04}.

The left panel of Figure 8 displays the relative difference in housing prices. Except for the very highest quality houses, the increased income inequality has lowered housing prices. The mean effect is -1.8%. That is, housing prices would be 1.8% higher in Helsinki, had the income inequality not increased as it did from 1998 to 2004. For houses near the middle of the distribution, the relative effect is about 3.5%. In absolute terms, the mean effect is about -3300 Euros.[9] The impact has been positive only at the top decile, and at most less than 1.5% in magnitude. For comparison, the increase in average real house prices was 6% per year in our data .

Similarly, we consider the effect of the overall income inequality by computing the equilibrium housing prices (again, given the same estimated quality distribution x^{98}) assuming that all households have the same income which equals the mean of θ^{04}. The right panel of Figure 8 displays the result. The impact of income inequality on the mean housing price is -22%, or 42000 euros.[10]

6 Conclusion

We have presented a new framework for studying the relationship between income distribution and housing price distribution. A key element of the model is that houses are heterogenous and indivisible. Importantly, the house quality distribution can take any form. That allows us to use the model to estimate the actual quality distribution in a fully non-parametric way.

Once we have the quality distribution, the model can be used to generate interesting counterfactuals. For instance, we can consider how housing price distribution would differ from the actual price distribution if the income distribution was different. As a first empirical application, we considered the impact of the recent increase in income inequality in Finland for the housing prices in the metropolitan area of Helsinki.

In terms of previous literature, our paper is placed between two quite separate strands. One,

[9]The dashed line in displays the results when the converse smoothing is used. The impact of increased inequality on mean house price is almost the same, -2.0%, but the distributional impact is more even.

[10]Under the converse smoothing strategy the corresponding impact is -12% or -26000 Euros

more empirically minded, in macroeconomics, uses dynamic equilibrium models and treats housing as a homogeneous asset, and another, mostly theoretical, in microeconomics that studies markets where at least one good or factor of production comes in heterogeneous indivisible units.

Appendix

Trading Lemma Consider two agents, b and s, with the same CES preferences but different endowments $\{x_b, \theta_b\}$ and $\{x_s, \theta_s\}$, where $x_s > x_b$. Trade involves a swapping of the indivisible goods (houses) and a payment in the continuous good $\Delta \in (0, \theta_b)$ by b, since she is endowed with the lower-quality house. In this sense b is the (potential) buyer. For trade to take place, there must exist a transfer Δ such that both agents are better off after the trade. Next we show that this is possible if and only if the buyer has higher autarky utility, i.e., if

$$\text{Buyer PC:} \qquad u(x_s, \theta_b - \Delta) > u(x_b, \theta_b), \tag{13}$$

$$\text{Seller PC:} \qquad u(x_b, \theta_s + \Delta) > u(x_s, \theta_s). \tag{14}$$

Inserting u from (3), and rearranging, these conditions can be expressed as

$$z > \theta_b^\rho - (\theta_b - \Delta)^\rho, \tag{15}$$

$$(\theta_s + \Delta)^\rho - \theta_s^\rho > z, \tag{16}$$

where $z \equiv \frac{\alpha}{1-\alpha}(x_s^\rho - x_b^\rho)$. Solving for Δ in the corresponding equalities gives buyer and seller reservation prices

$$\Delta_{\max} = \theta_b - (\theta_b^\rho - z)^{\frac{1}{\rho}}, \tag{17}$$

$$\Delta_{\min} = (\theta_s^\rho + z)^{\frac{1}{\rho}} - \theta_s. \tag{18}$$

There is trade if and only if $0 \leq \Delta_{\min} < \Delta_{\max} \leq \theta_b$. Define

$$\delta = u(x_b, \theta_b)^\rho - u(x_s, \theta_s)^\rho = \theta_b^\rho - \theta_s^\rho - z. \tag{19}$$

Note that δ is positive if i) buyer's autarky utility is higher than seller's and $\rho > 0$, or if ii) buyer's autarky utility is lower and $\rho < 0$. Solving for θ_b from (19) and substituting in (17), the trading condition $\Delta_{\min} < \Delta_{\max}$ becomes

$$(\theta_s^\rho + z)^{\frac{1}{\rho}} - \theta_s < (\theta_s^\rho + z + \delta)^{\frac{1}{\rho}} - (\theta_s^\rho + \delta)^{\frac{1}{\rho}}. \tag{20}$$

Note that z always has the opposite sign as ρ. The right side is equal to the left side if $\delta = 0$, but it is always strictly increasing in δ while the left side is constant. Hence trade is possible if and only if the household endowed with the better house has lower autarky utility.

References

ALONSO, WILLIAM (1964). *Location and Land Use.* Harvard University Press, Cambridge, MA.

BUITER, WILLIAM (2008). "Housing Wealth Isn't Wealth." *NBER Working Paper,* 14204.

CALOMIRIS, CHARLES; STANLEY D LONGHOFER AND WILLIAM MILES (2009). "The (Mythical) Housing Wealth Effect." *NBER Working Paper,* 15075.

EPPLE, 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.

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." *Working Paper.*

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*.

MIYAGAWA, EIICHI (2001). "House allocation with transfers." *Journal of Economic Theory*, 100(2), pp. 329–355.

MORETTI, ENRICO (2009). "Real Wage Inequality." *Working paper.*

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

459-85.

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 (20??). “Why Has House Price Dispersion Gone Up?” *Review of Economic Studies,* Forthcoming.

Figure 1.

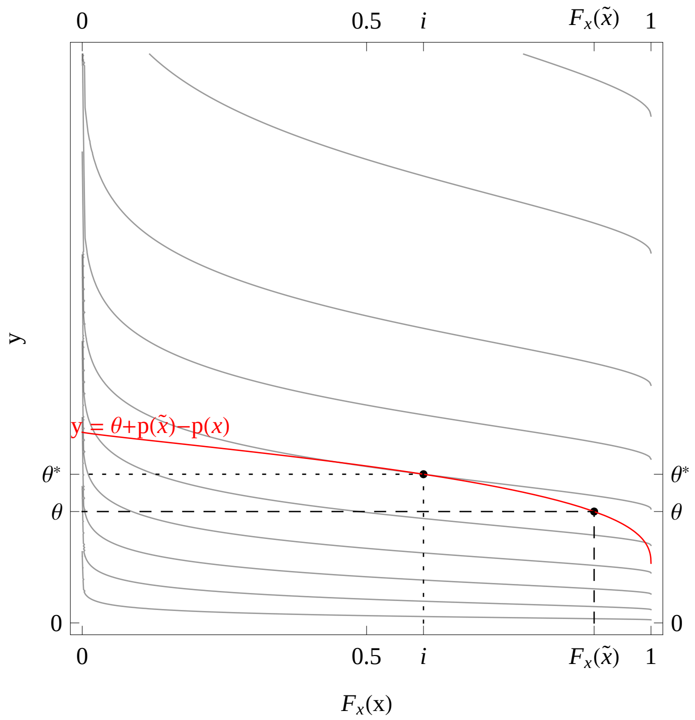
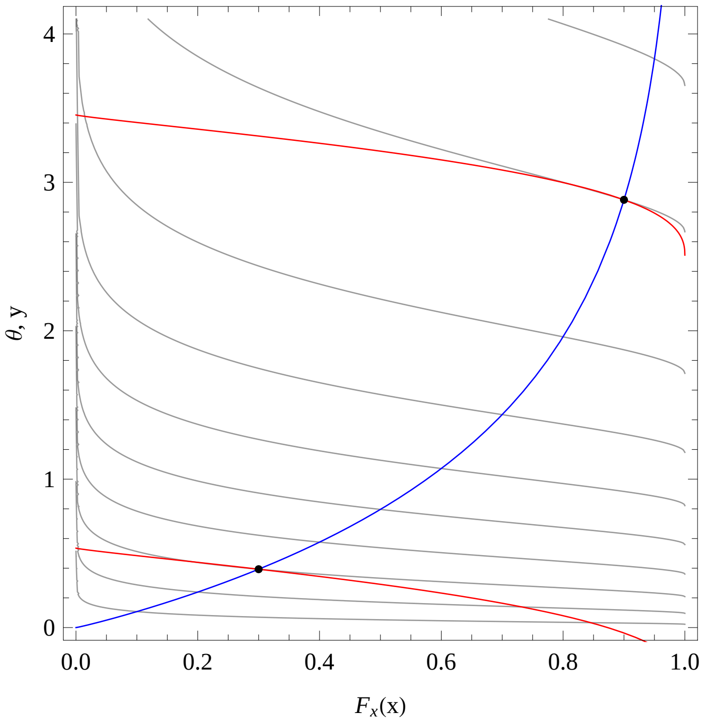

0
0.5
i
$F_x(\tilde{x})$
1
y
$y = \theta + p(\tilde{x}) - p(x)$
θ^*
θ
0
0
0.5
i
$F_x(\tilde{x})$
1
θ^*
θ
0
$F_x(x)$

Figure 2.



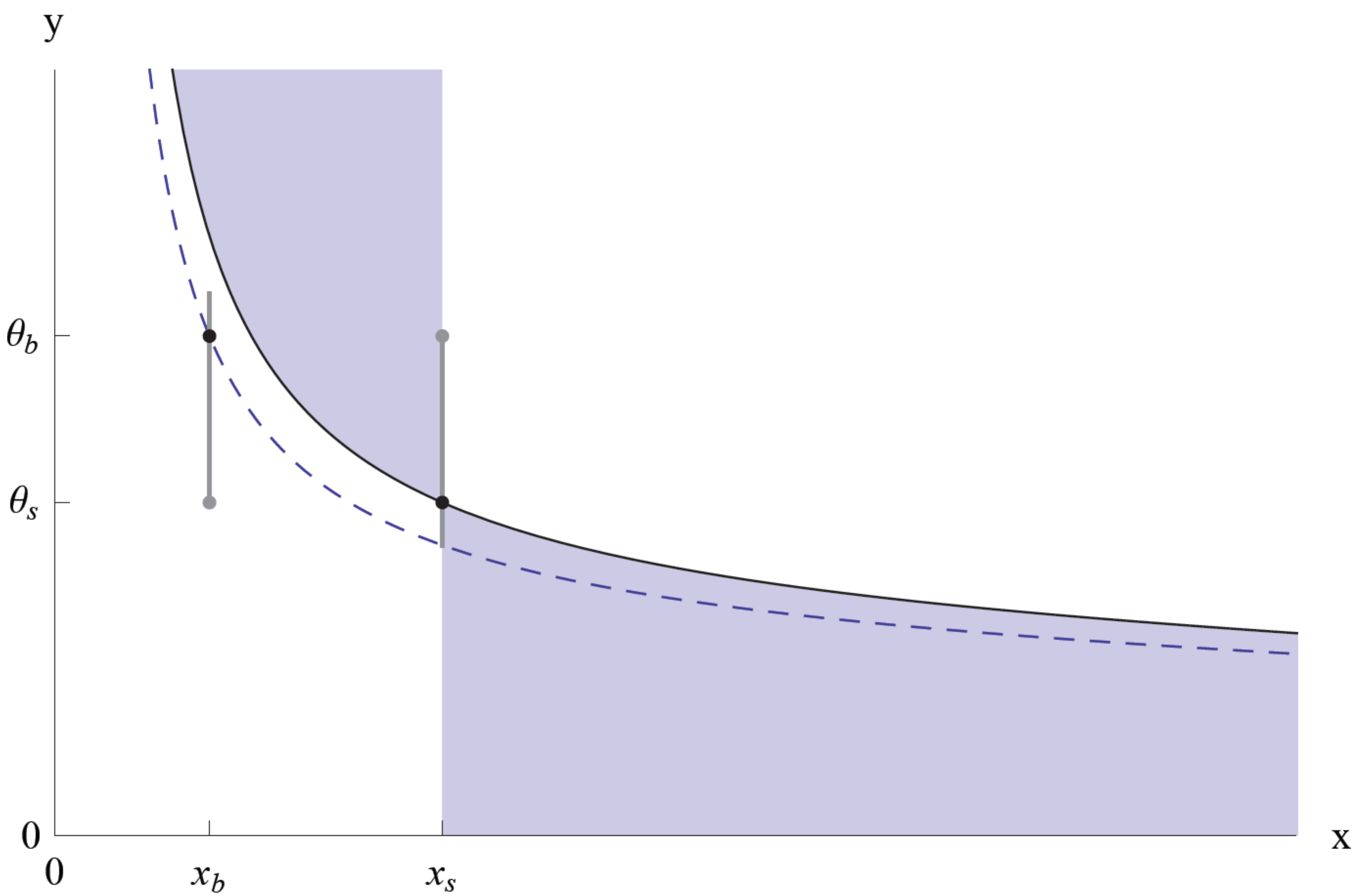


Figure 3. The endowments of a potential "seller" *s* and "buyer" *b* are depicted by the black dots. For mutually advantageous trade to be possible, the post-trade allocation must allow both households to move to a higher indifference curve. As houses *x* are indivisible, the new allocation must involve a pair-wise swap in the *x*-dimension, with *b* also transferring some money (composite good *y*) to *s*. Hence, after the trade, the consumption bundle of s will be on a vertical line that passes through $\{x_s,\theta_s\}$ and vice versa for *b*. Bundles that allow *b* to stay above her autarky indifference curve are marked with gray. Even the largest of such payments is not sufficient to bring *s* to the level of his autarky indifference curve.

Figure 4.

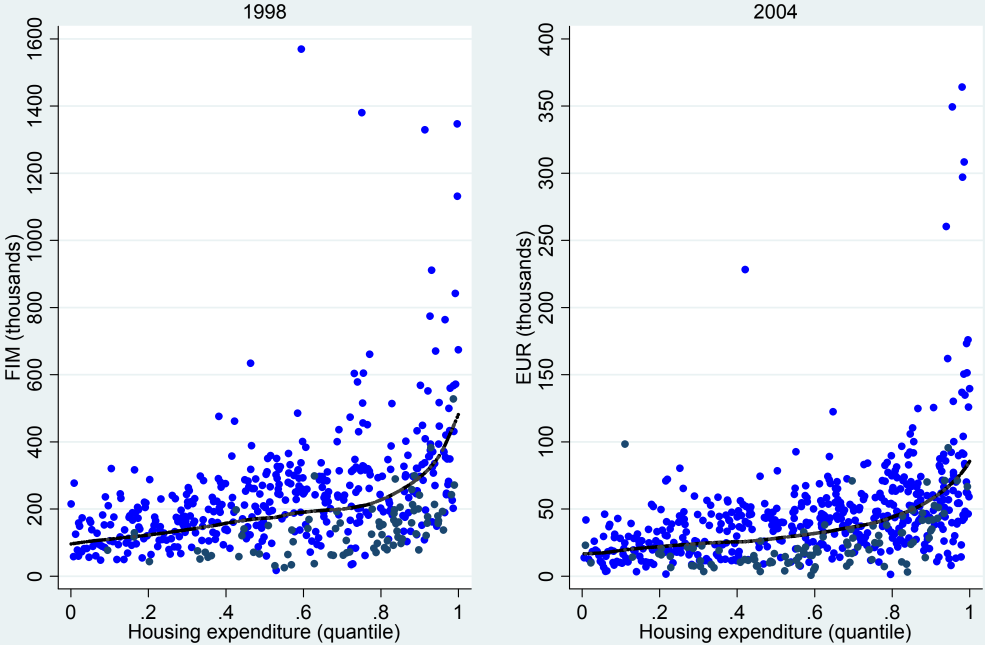

1998
2004
FIM (thousands)
EUR (thousands)
0
200
400
600
800
1000
1200
1400
1600
0
50
100
150
200
250
300
350
400
0
.2
.4
.6
.8
1
Housing expenditure (quantile)
Housing expenditure (quantile)

Figure 5

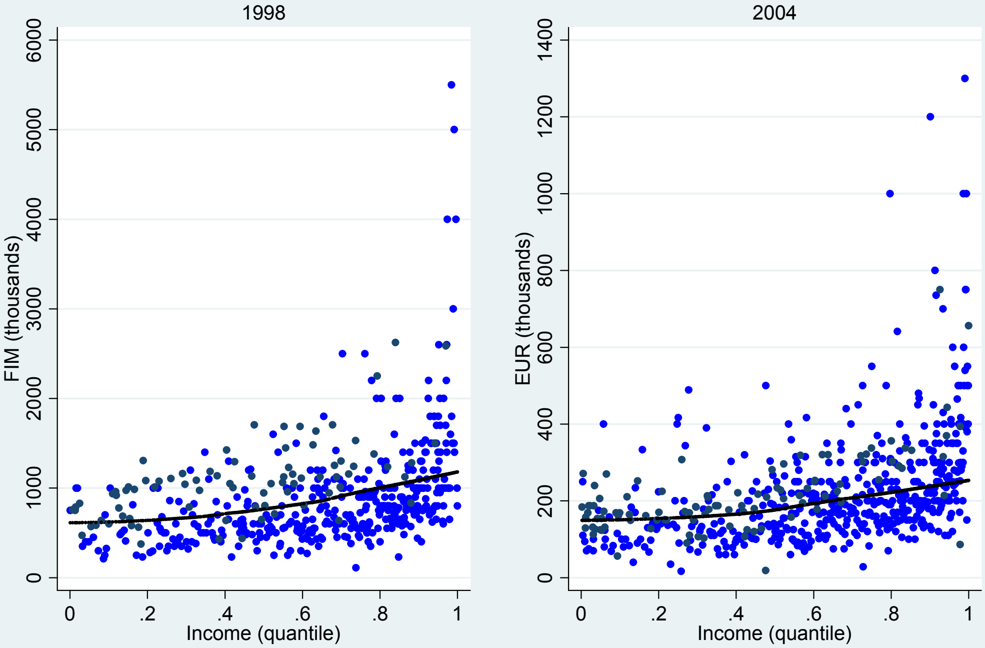
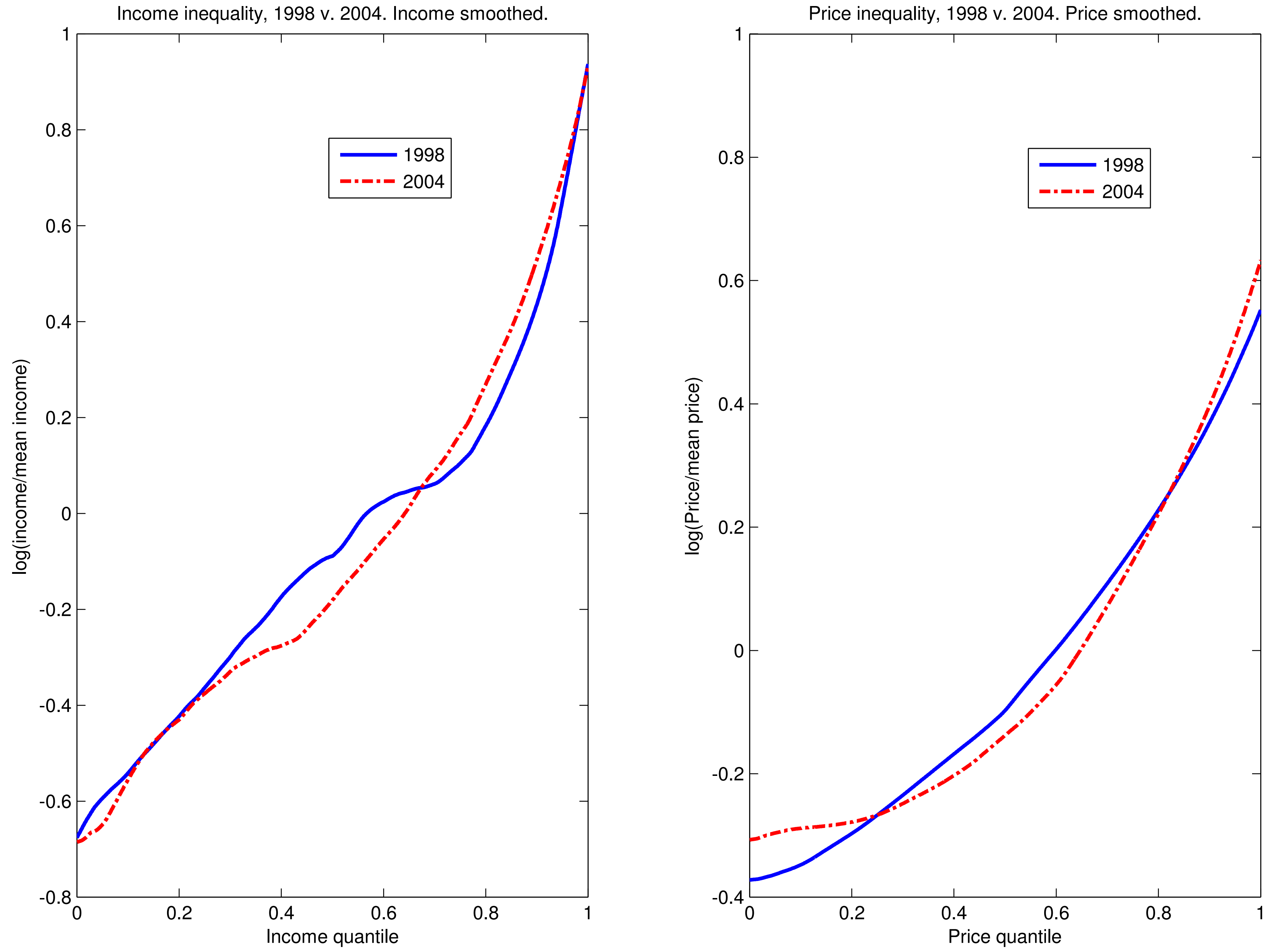
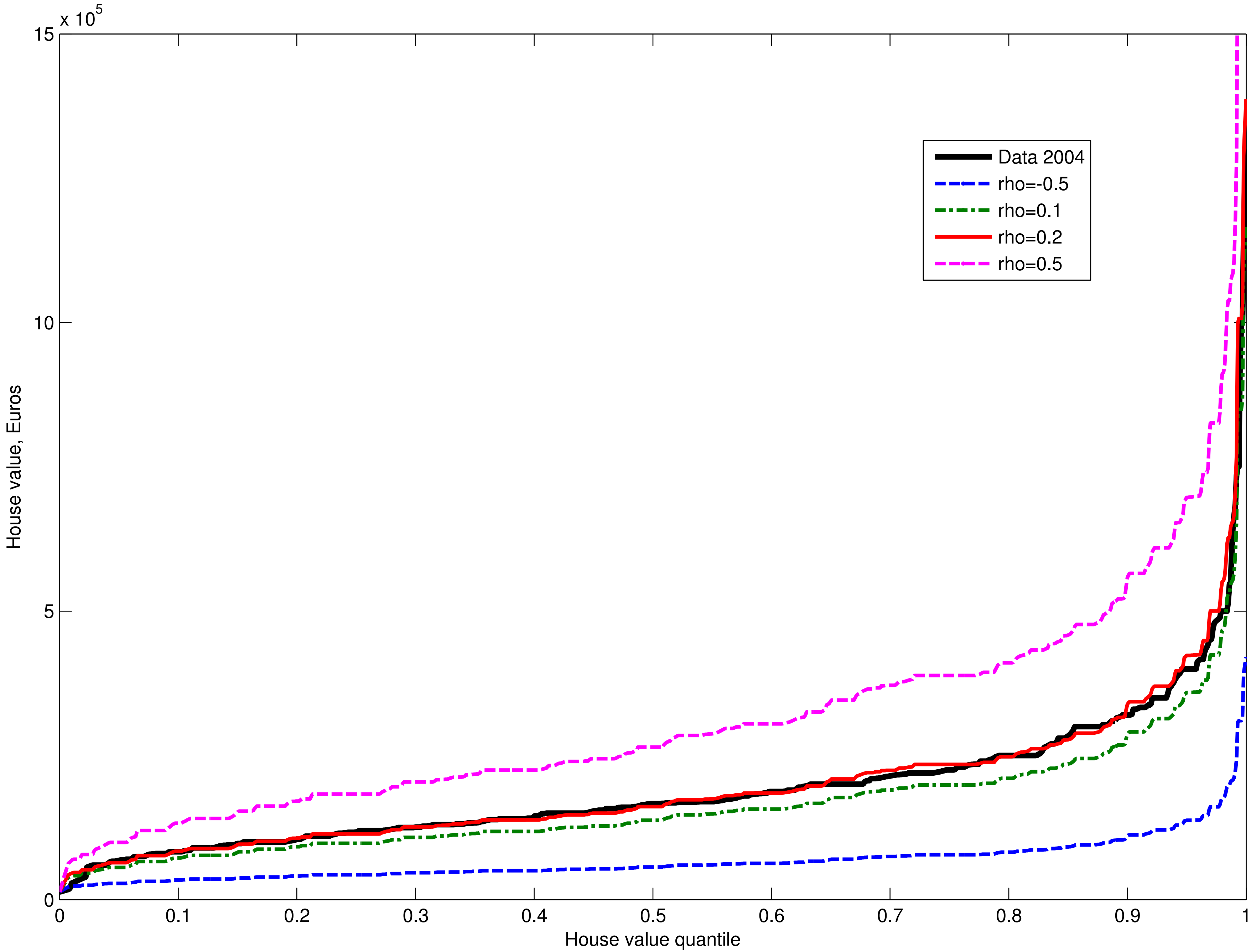

1998
FIM (thousands)
0
1000
2000
3000
4000
5000
6000
0
.2
.4
.6
.8
1
Income (quantile)
2004
EUR (thousands)
0
200
400
600
800
1000
1200
1400
0
.2
.4
.6
.8
1
Income (quantile)

Figure 6.



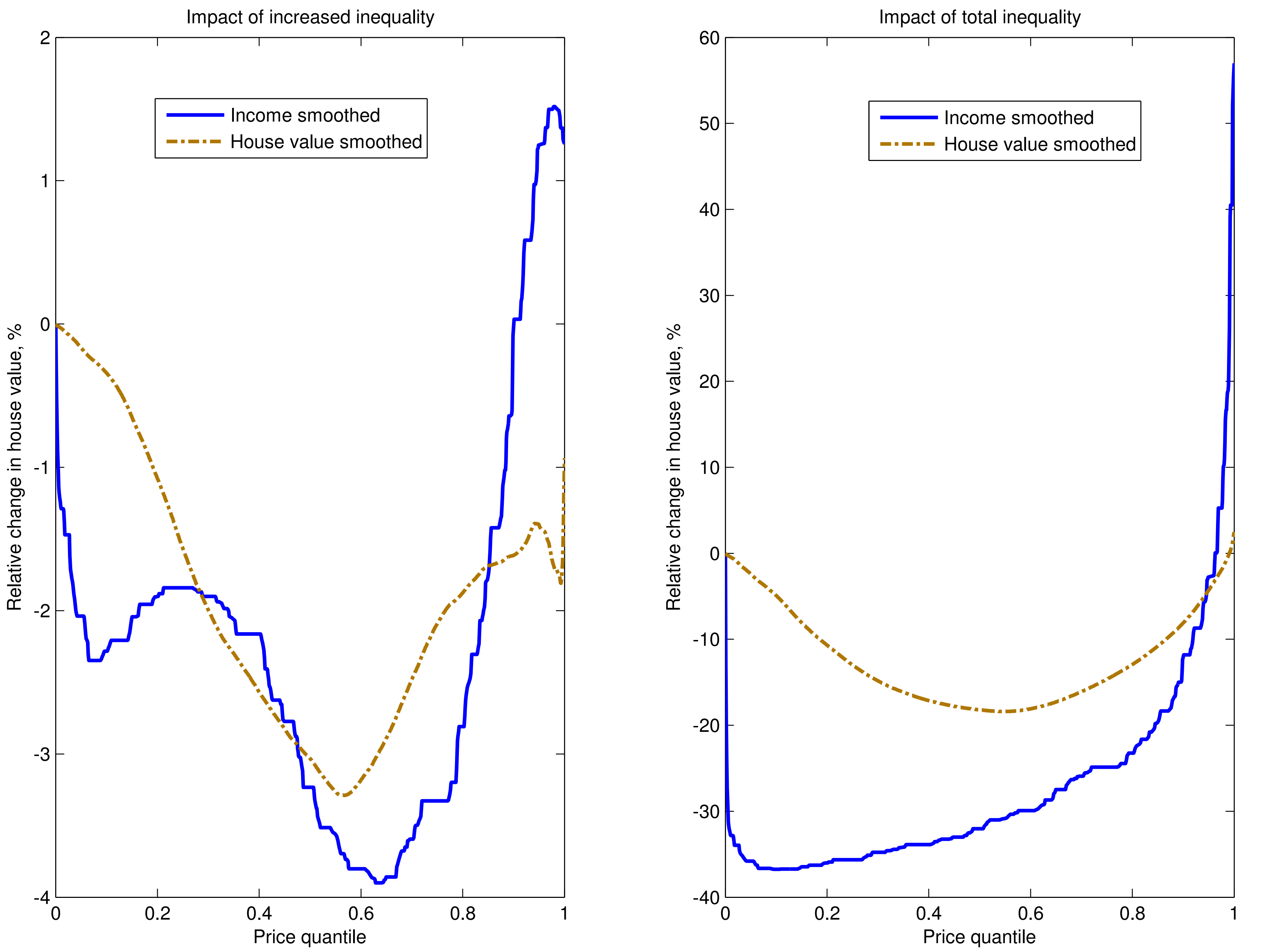
Income inequality, 1998 v. 2004. Income smoothed.
1998
2004
log(income/mean income)
Income quantile
1
0.8
0.6
0.4
0.2
0
-0.2
-0.4
-0.6
-0.8
0
0.2
0.4
0.6
0.8
1
Price inequality, 1998 v. 2004. Price smoothed.
1998
2004
log(Price/mean price)
Price quantile
1
0.8
0.6
0.4
0.2
0
-0.2
-0.4
0
0.2
0.4
0.6
0.8
1

Figure 7.



x 10^5
15
10
5
0
House value, Euros
0
0.1
0.2
0.3
0.4
0.5
0.6
0.7
0.8
0.9
1
House value quantile
Data 2004
rho=-0.5
rho=0.1
rho=0.2
rho=0.5

Figure 8.



Impact of increased inequality
Income smoothed
House value smoothed
2
1
0
-1
-2
-3
-4
Relative change in house value, %
0
0.2
0.4
0.6
0.8
1
Price quantile
Impact of total inequality
Income smoothed
House value smoothed
60
50
40
30
20
10
0
-10
-20
-30
-40
Relative change in house value, %
0
0.2
0.4
0.6
0.8
1
Price quantile