Income Distribution and Housing Prices: An Assignment Model Approach*

Niku Määttänen
HSE and HECER

Marko Terviö
HSE and HECER†

November 9, 2009

PRELIMINARY AND INCOMPLETE

Abstract

We present a framework for studying the relation between the distribution of income and the distribution of housing prices, based on an assignment model of households with heterogeneous incomes and houses of heterogeneous quality. We show how the unobserved distribution of quality can be inferred from the joint distribution of income and housing prices, and used to generate counterfactual price distributions under counterfactual income distributions. Using data from the Helsinki Metropolitan region, we find that the increase in income inequality from 1998 to 2004 caused the average price to be 1.8 to 3.2% lower than if income growth had been uniform across the population, while the impact on the top decile of housing prices was an increase of 0-6%.

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

†Helsinki School of Economics, and Helsinki Center for Economic Research. Arkadiankatu 7, 00101 Helsinki, Finland.

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.

In our view, 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 which is more or less fixed in short and medium run. Another key feature is that housing is a normal good, and takes up a large part of household expenditure. 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. Unlike in standard assignment models, we assume concave utility that gives rise to income effects. 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. We assume that households have the same preferences but are heterogeneous by income, so 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 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.

Our data is 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 1.8% to 3.2% higher, implying that the increase in inequality has resulted in lower house prices than uniform income growth. Only at the top decile has increased income inequality resulted in higher prices.

The reason why uniform income growth would have lead to higher prices at the bottom of the quality distribution is intuitive: as low-income households would have more income they would use some of it to bid for low-quality housing. 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 depends on how much the households at the respective part of the income distribution are willing to pay for the quality difference. Price decreases at the bottom of the distribution spill upwards in the distribution, partly undoing the local increase in willingness-to-pay among households whose incomes are now higher than under uniform growth. For this reason, while incomes have grown more than average in about 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 the ranking of households by income has stayed the same. In this case changes in housing wealth really aren't changes in wealth from anyone's point of view.[1] More generally, if households' positions in the income distribution

[1] See also "Housing Wealth Isn't Wealth" by Buiter 2008. In our setup, only increases in house quality could

change then they 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 effect of the housing prices is that the newly rich lose out and the no-longer quite so rich gain relative to a "naive" measure where income were equated with consumption.

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 assignment model with concave utility; for a review of assignment models see Sattinger (1993). These are models of matching markets with symmetric information and no frictions. 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. When combined with assortative matching, assignment models yield unique 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 positive assortative matching arises from a productive complementarity between job types and worker types, Sattinger (1979) and Teulings (1995), or between workers themselves in a team production setting, as in Kremer (1993). Our nonparamteric method for inferring the unobserved type distribution from observed price distributions was introduced in Terviö (2008). In our current paper there is no complementarity, but equilibrium nevertheless involves assortative matching by income and house quality due to housing being a normal good.

Matching models have long been applied to the housing market from a more theoretical

result in aggregate increases in (reasonably defined) housing wealth.

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 knowledge there are no papers on matching markets where i) there are both indivisible and continuous goods and ii) 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 that results in what 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 land 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.

Van Nieuwerburgh and Weill (2006) build a dynamic model with housing to explain the increase in housing price dispersion across US cities. In their model, there is matching between individual ability and regional productivity.

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 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. Households have the same preferences but different endowments. Every household is endowed with and consumes exactly one house. Household income 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 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 is described by a point in $[0, 1] \times \mathbb{R}_+$, where the horizontal dimension represents 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.

The composite good y is used as the numeraire, while p is the equilibrium price function for house quality. Budget constraints are downward sloping curves, as 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 worst occupied unit of housing, of quality $x_0 > 0$, is also available as an outside option at an exogenous price $p_0 \geq 0$. The joint distribution of endowments (x, θ) is continuous and without gaps in $[x_0, x_1] \times \mathbb{R}_+$, 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 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 $\{x_0, 0\}$, 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 that whenever the marginal rates of substitution differ between the two households—which, generically speaking, would always be 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, since 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 so that both households could be better off than after swapping houses; see Figure 2b for the illustration. For CES utility, trade is possible if and only if the household with the better house has lower autarky utility. [Appendix proof: to do] The point is that each endowment comes with a trading set, which includes all those other endowments with which mutually profitable trade is possible. In equilibrium, it must be the case that no point on the no-trade curve can be 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, and that utility is increasing along it, although consumption y may not be.

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 buys the better and more expensive house. The twist here is that wealth is endogenous, because the value of the endowment depends on p. Therefore the equilibrium matching is not obvious (unless p is known) 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 for $\{x(i), \theta^*(i)\}$ is equal to i, i.e., to the mass of houses of quality $x(i)$ or less.[2] In general, this leads to a complicated partial differential equation. However, by discretizing the house types, the equilibrium can be solved numerically using standard recursive methods.

Absentee landlord model

Now consider an otherwise similar economy, but where all houses are 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 different 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

[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).

result in every household buying a house of the same quality rank as is their rank in the wealth distribution, so that

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

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:

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

This is the key equation of the model. Combined with the exogenous boundary condition $p\left(0\right) = p_0$ 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 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 quality difference. This depends on their marginal rate of substitution between house quality and other goods, which in general depends 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.

It is worth noting that, in the light of our model, the claim that an increase in income inequality should lead to an increase in the prices of best houses (land rents of most desirable locations) is incorrect. To see this, suppose some wealth is redistributed from poor to rich, holding average wealth fixed. It is true that this will increase the local "price gradient" (2) at the top quantiles, as the willingess-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 prices, 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 started at some arbitrary continuously distributed endowment, and then traded to equilibrium. As we saw above, the new allocations should be located on a no-trade curve, where there is perfect rank correlation between house quality and total wealth. Now (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 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 the income level θ, which is a fixed characteristic of the dynasty. As we will show later [to do !!], this model has a unique steady state and it involves 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).

3.1 The case with CES

For the empirical application we assume CES utility,

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

The equilibrium price condition (2) is now

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

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 and Cobb-Douglas utility Under some specific assumptions, the model equilibrium prices have a closed-form solution. Our empirical applications do not require such solution, but they would be desirable in theoretical extensions. Here we present an example of a closed-form solution.

Assume that both wealth and house quality follow Pareto distributions, so that w˜Pareto(w_0, η) and x˜Pareto(x_0, γ), where $p_0 \in (0, w_0)$ and $\eta, \gamma > 2$. Then the quantile functions are $w(i) = w_0 (1-i)^{-\frac{1}{\eta}}$ and $x(i) = x_0 (1-i)^{-\frac{1}{\gamma}}$. If utility takes the Cobb-Douglas form, so that $\rho \to 0$ in (3), then (4) has a closed form solution:

$$p(i) = w_0 (1-i)^{-\frac{1}{\eta}} - (w_0 - p_0)(1-i)^{\frac{\alpha}{\gamma(1-\alpha)}} - \frac{(1-\alpha)\gamma w_0}{\alpha\eta + (1-\alpha)\gamma} \left((1-i)^{-\frac{1}{\eta}} - (1-i)^{\frac{\alpha}{\gamma(1-\alpha)}} \right). \tag{6}$$

The expenditure share of housing $a(i) = p(i)/w(i)$ has the limit

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

The expenditure share a is everywhere increasing (decreasing) if $a(1) > (<) \, p_0/w_0$. 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 is strictly increasing 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 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. This would mean that, at the margin, a household could always sell or buy any fraction of their house at the same unit price.

3.2 Discussion

Suppose all households have the same wealth $\bar{w} > 0$ and preferences are Cobb-Douglas. Then (4) simplifies to

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

Here prices must make every household indifferent between every housing unit. It is easy to see why the expenditure share of housing cannot be constant for the population even if income and preferences are the same for all – some households must end up with the lower quality houses and for this they must be compensated with higher consumption on other goods.

More generally, the expenditure share of housing depends on the shapes of the income and house quality distributions. Even under Cobb-Douglas utility, the expenditure share of housing is not directly given by α because the prices faced by the consumers are nonlinear. For the same reason, the level of consumption of the composite good need not be monotonic in wealth. This is easy to understand when there is a steep increase in the quality of housing at some point of the distribution – clearly the price of housing must jump at that point, even if income is smoothly distributed. This in turn implies a negative relation between wealth and consumption y.

4 Inference

Assuming a utility function, the model can be used to infer the unobserved distribution of x from the observed relation between household income and house prices. 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. $x_0 = 1$. Later we make sure to only use the inferred distribution of x to answer questions where the constant x_0 washes out from the answers. A number of interesting counterfactuals 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. 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 the equilibrium conditions for prices are

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

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 $Y\left(u\left(x, y\right), x\right) = y$, 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

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

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

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

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 monetary variables. This is because, in the utility function, changing α is equivalent to changing the units of x. But 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.

4.1 Constructing 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 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 actual value.

We can study the impact of hypothetical demand shocks, supply shocks, and double-sided shocks...

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. Indeed, according to OECD (2008, figure 1.2), Finland had the largest increase in income inequality from mid-1990s to mid-2000s of all OECD countries.

5.1 Data and smoothing

Our data is the 1998 and 2004 Wealth Surveys from the Statistics Finland. These are the two most recent Wealth Surveys [but 2008 should be available soon]. The Survey has detailed portfolio information from about 2500 households. We focus on homeowners in the Metropolitan area of Helsinki.[4] That leaves us 402 households in the 1998 survey and 481 households in the 2004 survey.

We use just two variables. The first is household's total monetary income during the last year. This measure includes transfers and taxes. We take the total monetary income as our proxy for permanent income. 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.

For both 1998 and 2004, we observe the joint distribution of θ and p. Since we use the "post-

[4]Almost 40% of the households in the survey are renters. In principle, there is no problem in including rental housing into the model, at least if we think of the landlords as being absent from the model. However, most of the rental dwellings are part of social housing and the tenants are selected on the basis of social and financial needs. The rents of these houses are regulated and do not reflect well the market value. To our understanding, there has not been major changes in the supply of social housing or in the means testing criteria to social housing from 1998 to 2004.

trade" version of our model in this empirical application, we need to estimate a relationship that reduces θ and p to a common increasing order. Of course, in the data the two variables are not perfectly rank correlated.

One way to smooth the data is to non-parametrically estimate $p\,(i)$ as $E[p|F_{\theta}\,(\theta)\;=\;i]$. Alternatively, we can estimate $\theta(i)$ as $E[\theta|F_p\,(p)\;=\;i]$. We prefer the latter option because the observed income θ is likely to be a noisy proxy for permanent income, while the reported house value is presumably a relatively accurate measure of the true market value. Therefore, we take the latter option as our benchmark case but we also experiment with the first option. In both cases, we get income and price distributions for 1998 and 2004 that are perfectly rank correlated. We denote these smoothed distributions by $\{\theta^{98}, p^{98}\}$ and $\{\theta^{04}, p^{04}\}$.

Figure 3 displays the joint distributions together with the kernel regressed relationships $E[\theta|F_p\,(p) = i]$. Clearly, there is a strong positive relationship between the rank of the house value and household's income. Figure 4 in turn displays the kernel regressed relationships $E[p|F_{\theta}\,(\theta) = i]$.

Figure 5 displays the distributions of log income relative to mean log income in 1998 and 2004. Clearly, there is a substantial increase in income inequality from 1998 to 2004. It is this change in the income distribution, that allows us to infer preferences. Figure 6 displays the distribution of log housing prices relative to mean log housing prices in 1998 and 2004. 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 $w^{98}(i) = \theta^{98}(i) + rp^{98}(i)$. 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 value. Alternatively and equivalently, we could write $w^{98}(i) = \theta^{98}(i)/r + p^{98}(i)$. We fix the interest rate at $r = 0.05$. Changing r over a reasonable range $(3 - 7\%)$ does not substantially change our inferred elasticity parameter ρ or the results from the counterfactual experiments that we present below.[5]

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. Here it should be noted that we set the price of the lowest quality house, $p(0)$, equal to $p^{04}(0)$. As explained above, the model does not explain the level of housing prices but the difference to the price of the lowest quality house.

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. Visual inspection suggests that $\rho \approx 0.3$ provides the best fit.

We also inferred ρ estimating $p(i)$ as $E[p|F_\theta(\theta) = i]$. Figure 8 displays the empirical and predicted price distributions for different values of ρ. The best fit is again provided by ρ around 0.3. In other words, in terms of inferring preferences, it does not seem to make a big difference whether we smooth the data by estimating income as a function of house value quantile or vice versa.

5.3 Counterfactuals

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 $\rho = 0.3$ and given the inferred quality distribution, x^{98}, we compute the equilibrium housing price distribution that would be

[5] It also seems that the inferred elasticity parameter does not change much even if we assume that the interest rate is different in 1998 and in 2004. Essentially, 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.

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, θ^{04}.

Figure 9 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 -3.20%. That is, housing prices would be 3.2% higher in Helsinki, had the income inequality not increased as it did from 1998 to 2004. For houses in the bottom of the distribution, the relative effect is almost 10%. In absolute terms, the mean effect is about -4200 euros.

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}. Figure 11 displays the result. The impact of income inequality on the mean housing price is -23.2%.

Figures 10 and 12 relate to the same experiments but with data smoothed by estimating $p(i)$ as $E[p|F_\theta(\theta) = i]$. The effects are quite similar to those shown in figures 9 and 11. However, the impacts on mean housing price are somewhat smaller now: The impact of the increased income inequality is now 1.75% and the impact of the overall income inequality is 13.7% on the mean housing price.

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, 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.

This research is obviously work in progress. One topic that we plan to study with our framework is supply shocks. For instance, because of a move of a container harbour away from its current location close to Helsinki city center, a large area of land in a very attractive location has just recently become available for residential construction. This will be a major housing supply shock in Helsinki. If the quality of new housing can be placed in the current quality distribution, then our model can be used to predict the effect of this supply shock on the housing price distribution. We believe that our model would also be an interesting set-up to consider the distributional impacts of various tax-and-transfer schemes, particularly those that are directly related to housing.

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.

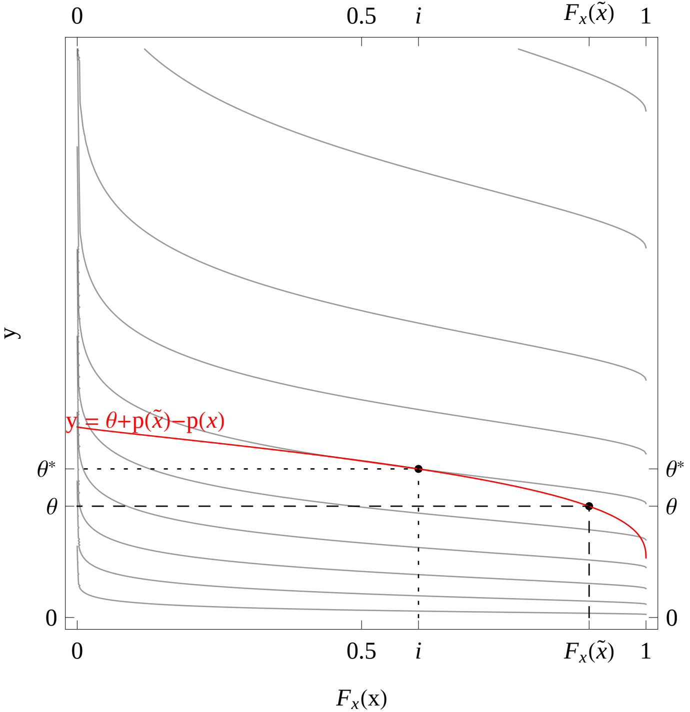
TERVIÖ, MARKO (2008). “The Difference that CEOs Make: An Assignment Model Approach.” *American Economic Review* 98(3), pp. 642–668.

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 (2006). “Why Has House Price Dispersion Gone Up?” *NBER Working Paper*, 12538.

Figure 1.



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

Figure 2.

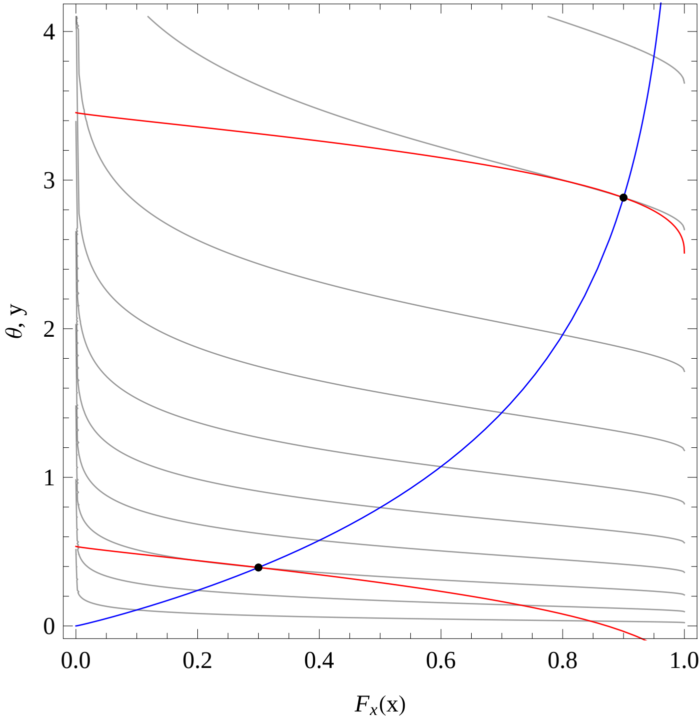
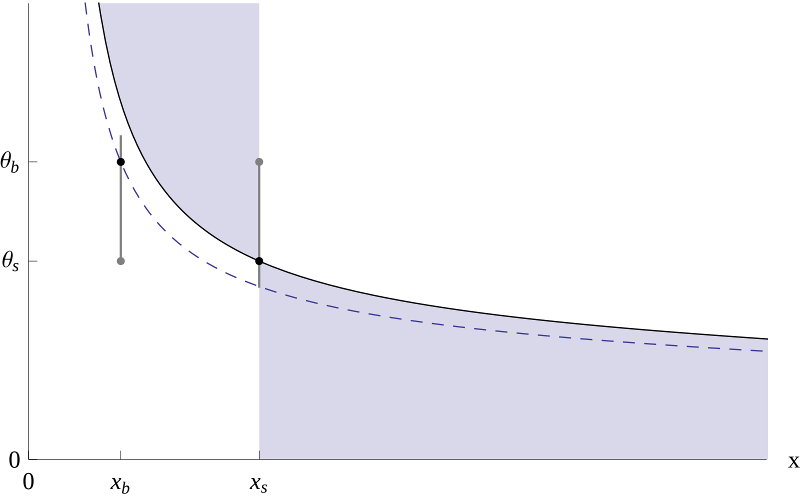


Figure 2b.



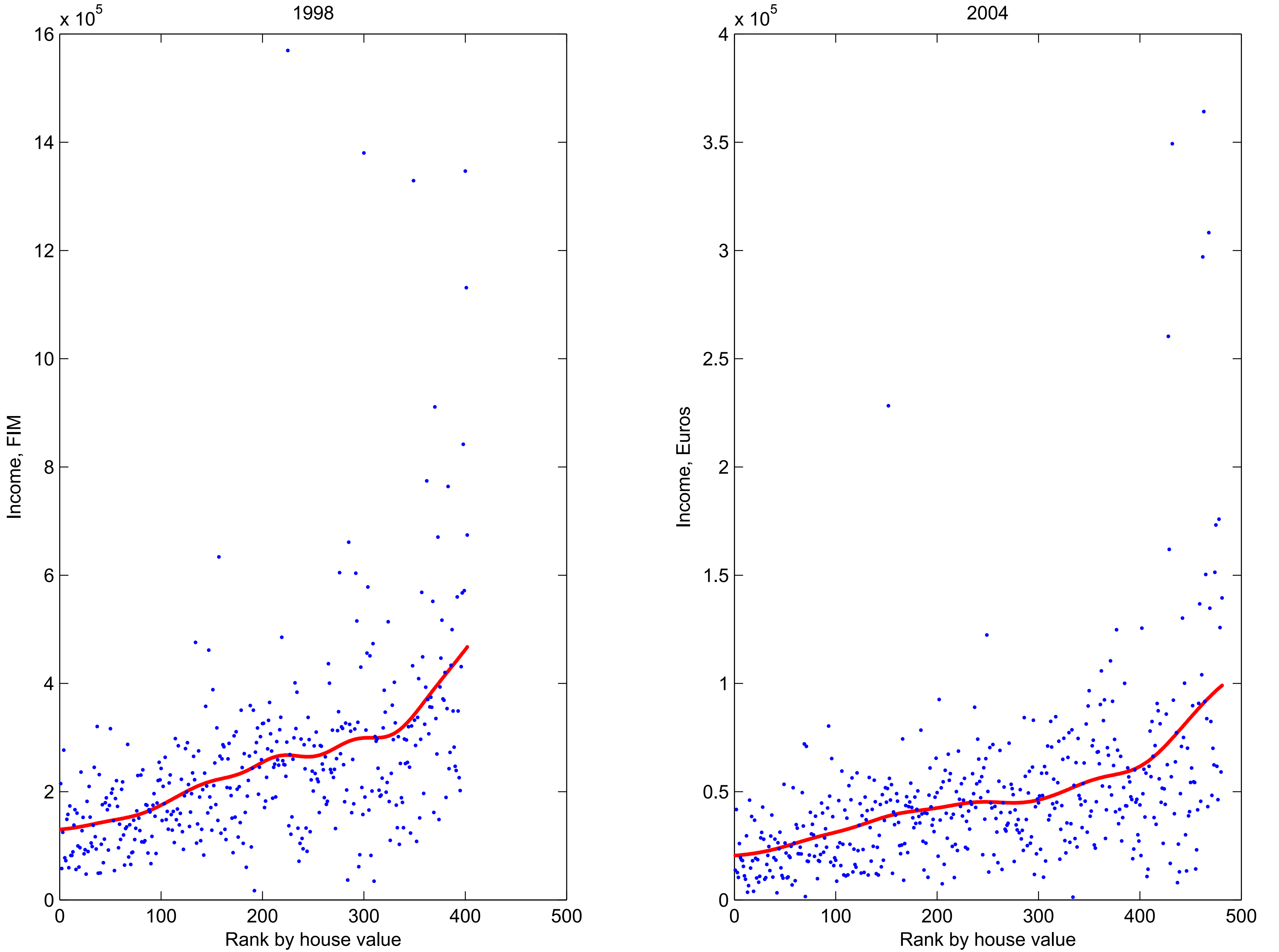
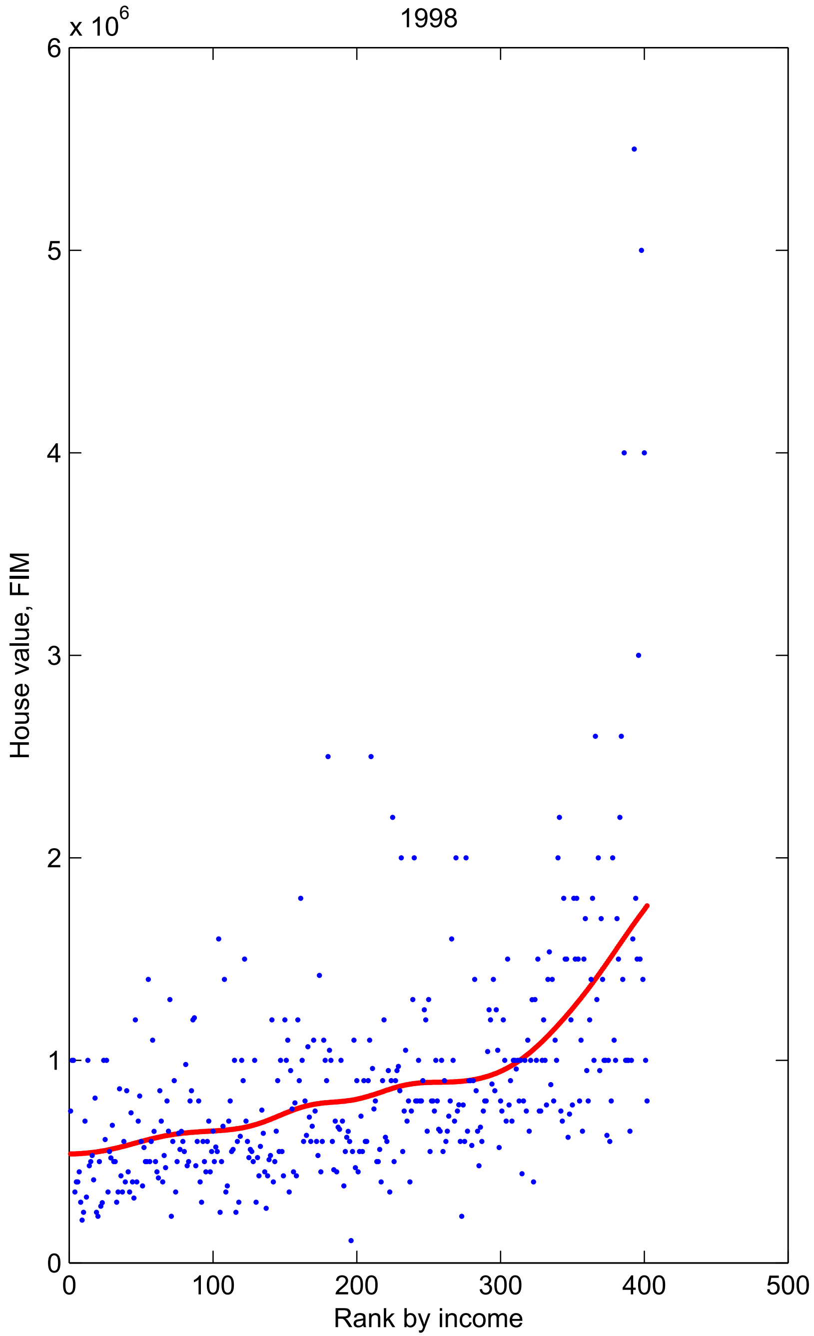
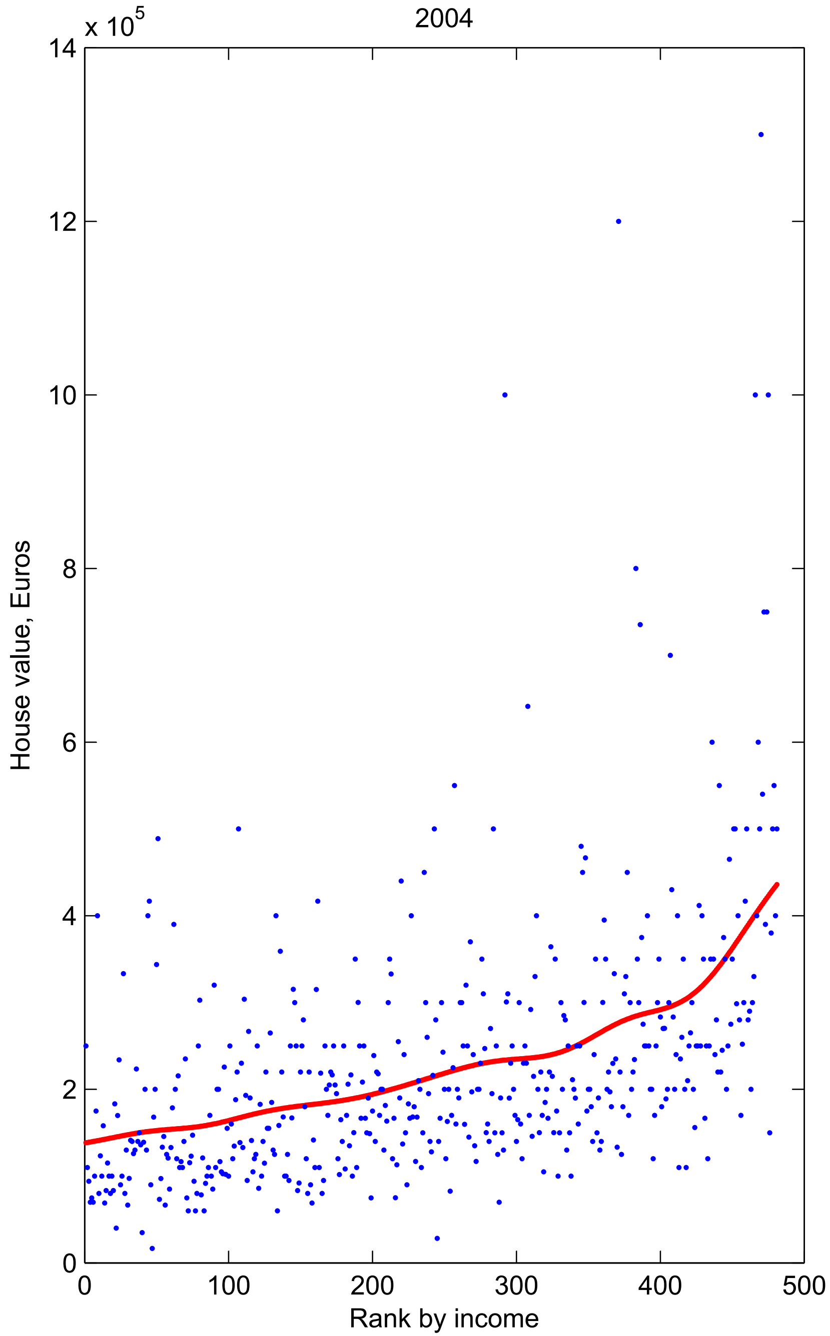

1998
x 10^5
Income, FIM
Rank by house value
0
2
4
6
8
10
12
14
16
0
100
200
300
400
500
2004
x 10^5
Income, Euros
0
0.5
1
1.5
2
2.5
3
3.5
4

Figure 3



1998
x 10^6
House value, FIM
Rank by income
0
1
2
3
4
5
6
0
100
200
300
400
500



2004
x 10^5
House value, Euros
Rank by income
0
2
4
6
8
10
12
14
0
100
200
300
400
500

Figure 4

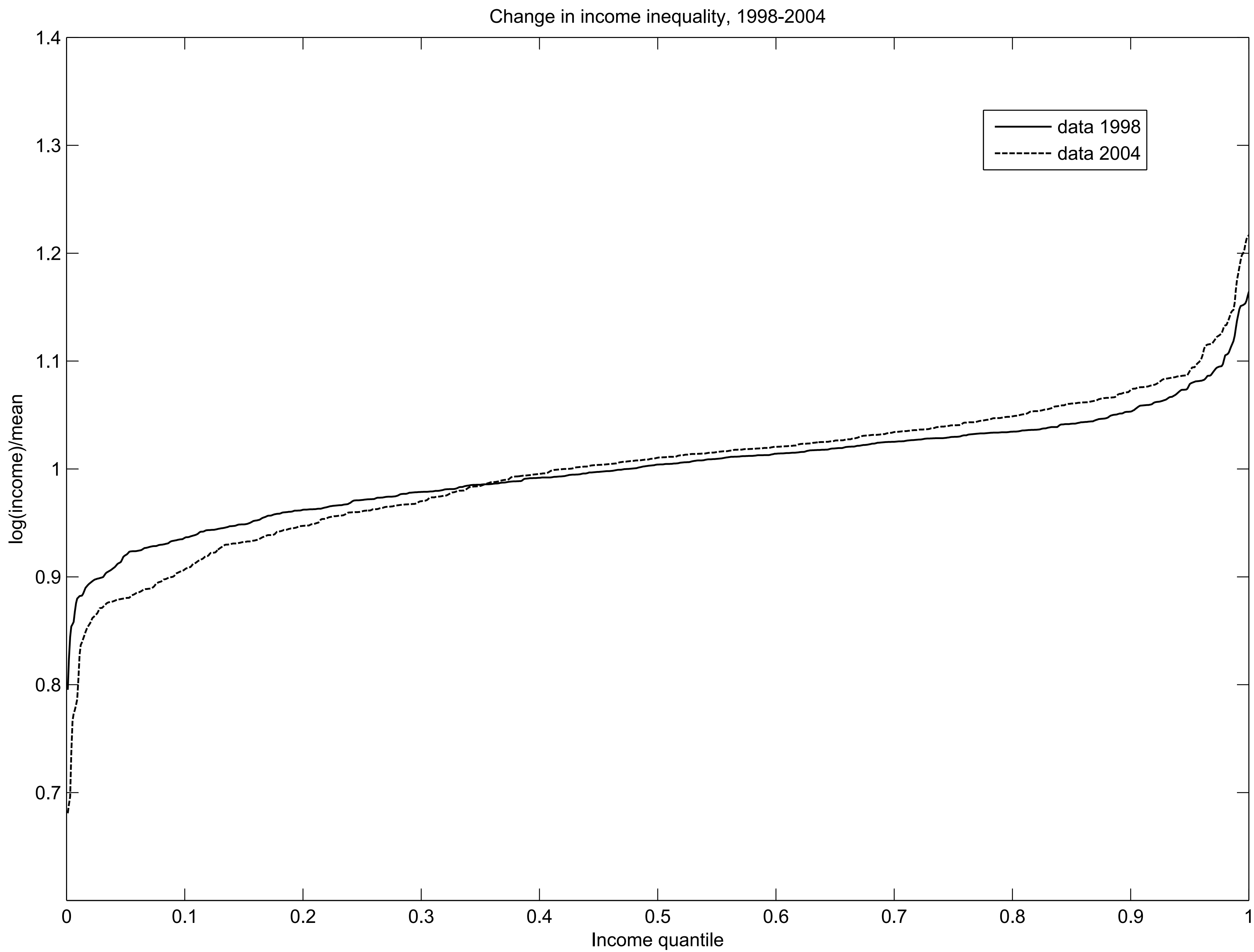

Change in income inequality, 1998-2004
data 1998
data 2004
log(income)/mean
Income quantile
1.4
1.3
1.2
1.1
1
0.9
0.8
0.7
0
0.1
0.2
0.3
0.4
0.5
0.6
0.7
0.8
0.9
1

Figure 5

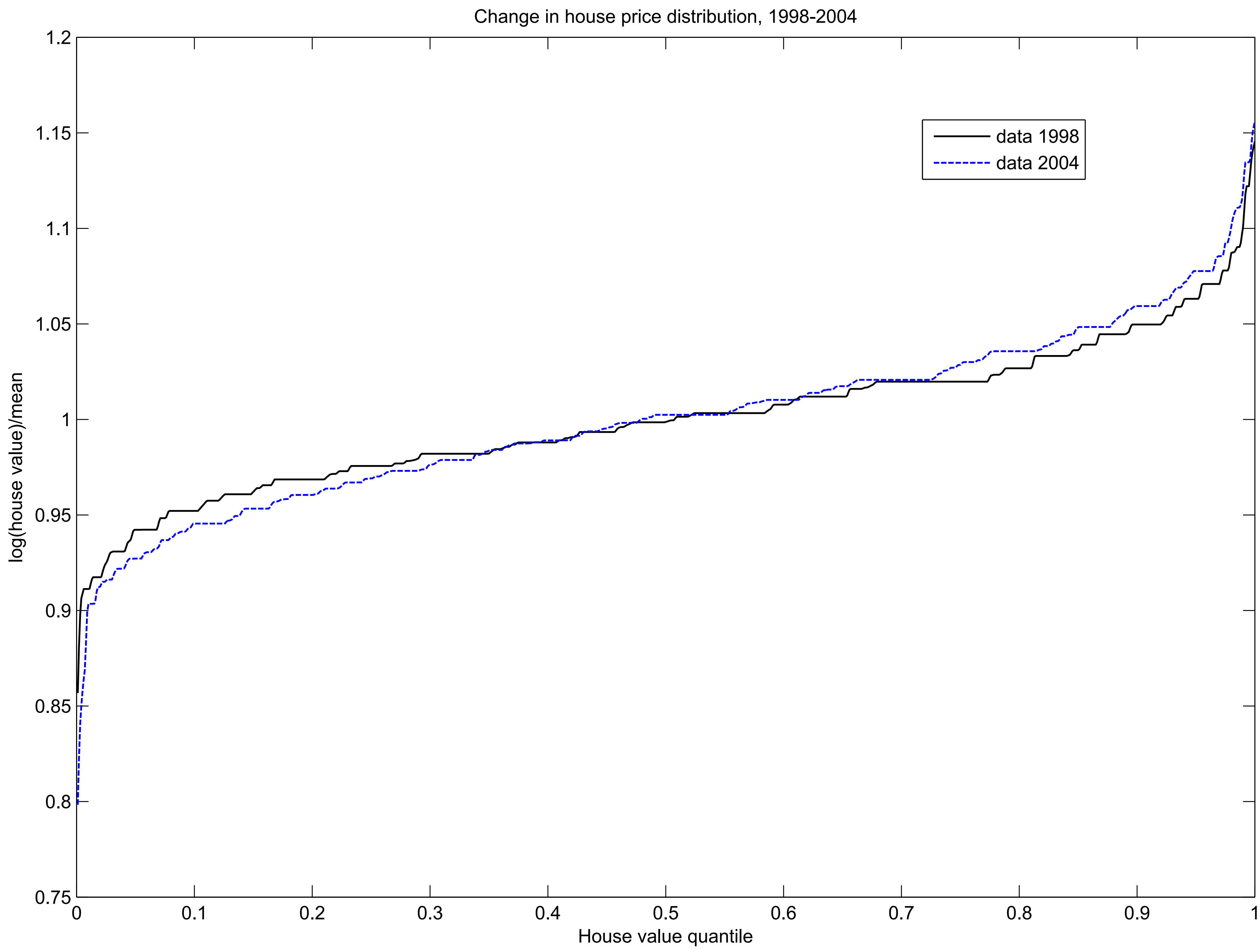
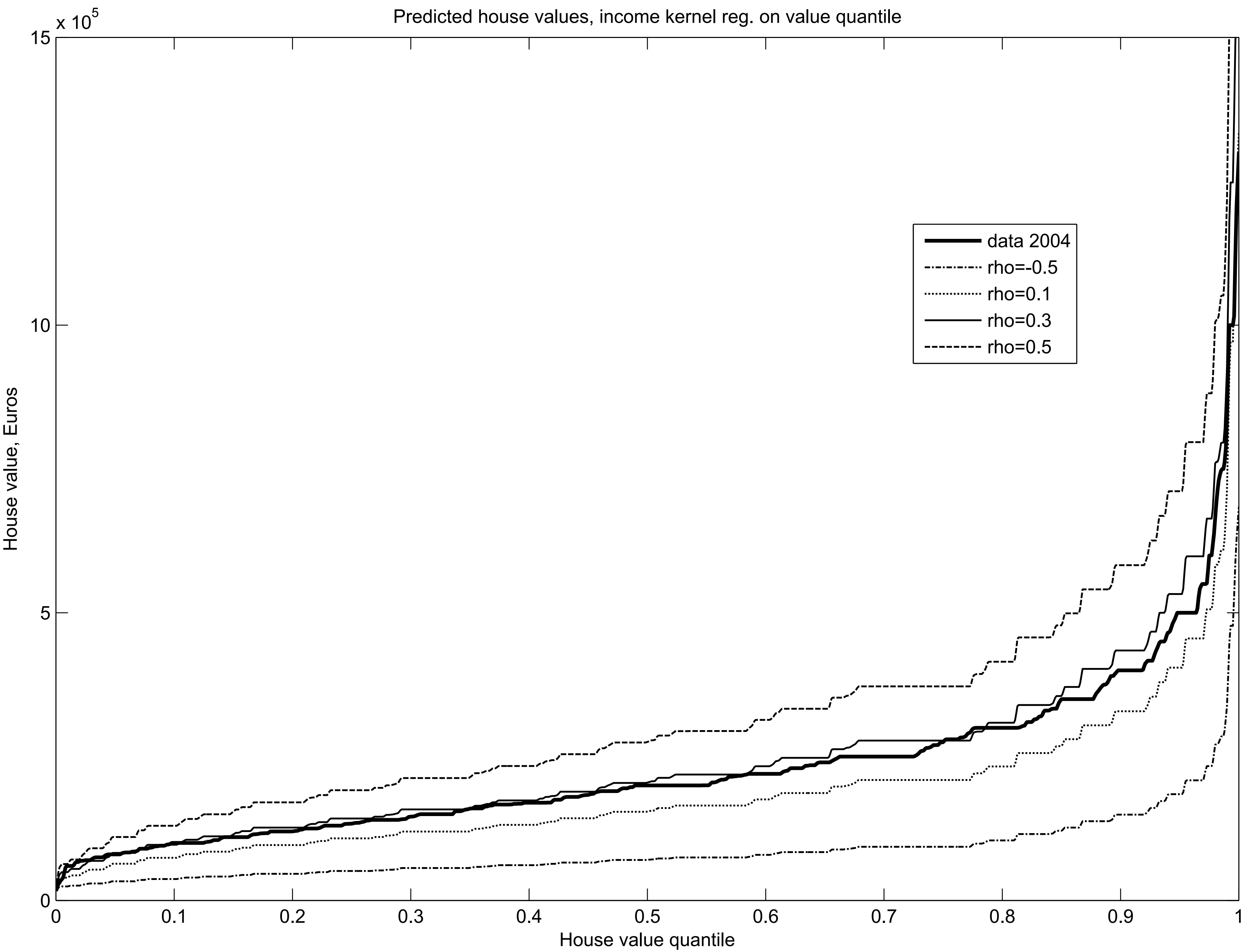
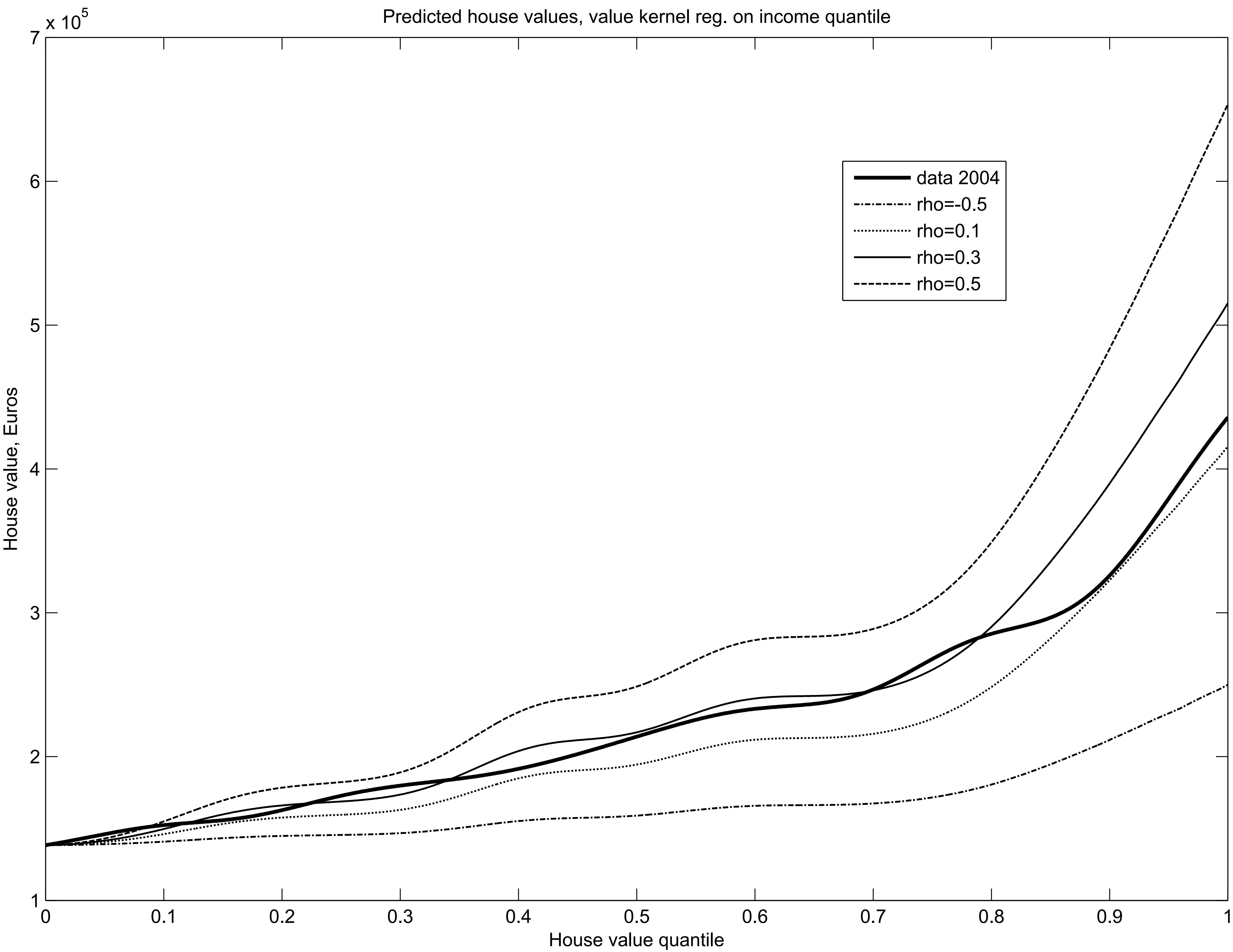

Change in house price distribution, 1998-2004
data 1998
data 2004
log(house value)/mean
House value quantile
1.2
1.15
1.1
1.05
1
0.95
0.9
0.85
0.8
0.75
0
0.1
0.2
0.3
0.4
0.5
0.6
0.7
0.8
0.9
1

Figure 6



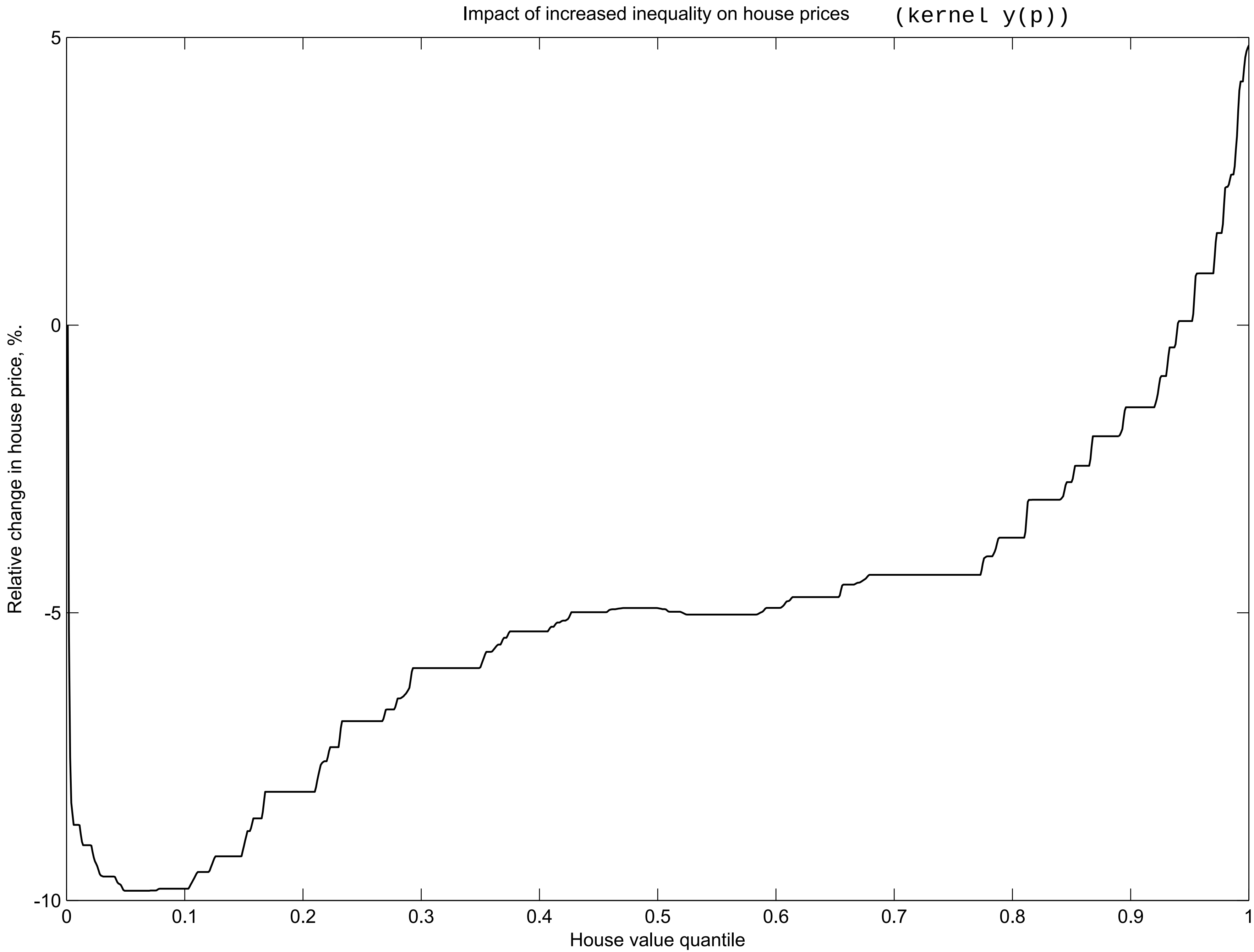
Predicted house values, income kernel reg. on value quantile
x 10^5
15
10
5
0
House value, Euros
data 2004
rho=-0.5
rho=0.1
rho=0.3
rho=0.5
0
0.1
0.2
0.3
0.4
0.5
0.6
0.7
0.8
0.9
1
House value quantile

Figure 7



Predicted house values, value kernel reg. on income quantile
x 10^5
7
6
5
4
3
2
1
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=0.3
rho=0.5

Figure 8



Impact of increased inequality on house prices (kernel y(p))
5
0
-5
-10
Relative change in house price, %.
0
0.1
0.2
0.3
0.4
0.5
0.6
0.7
0.8
0.9
1
House value quantile

Figure 9

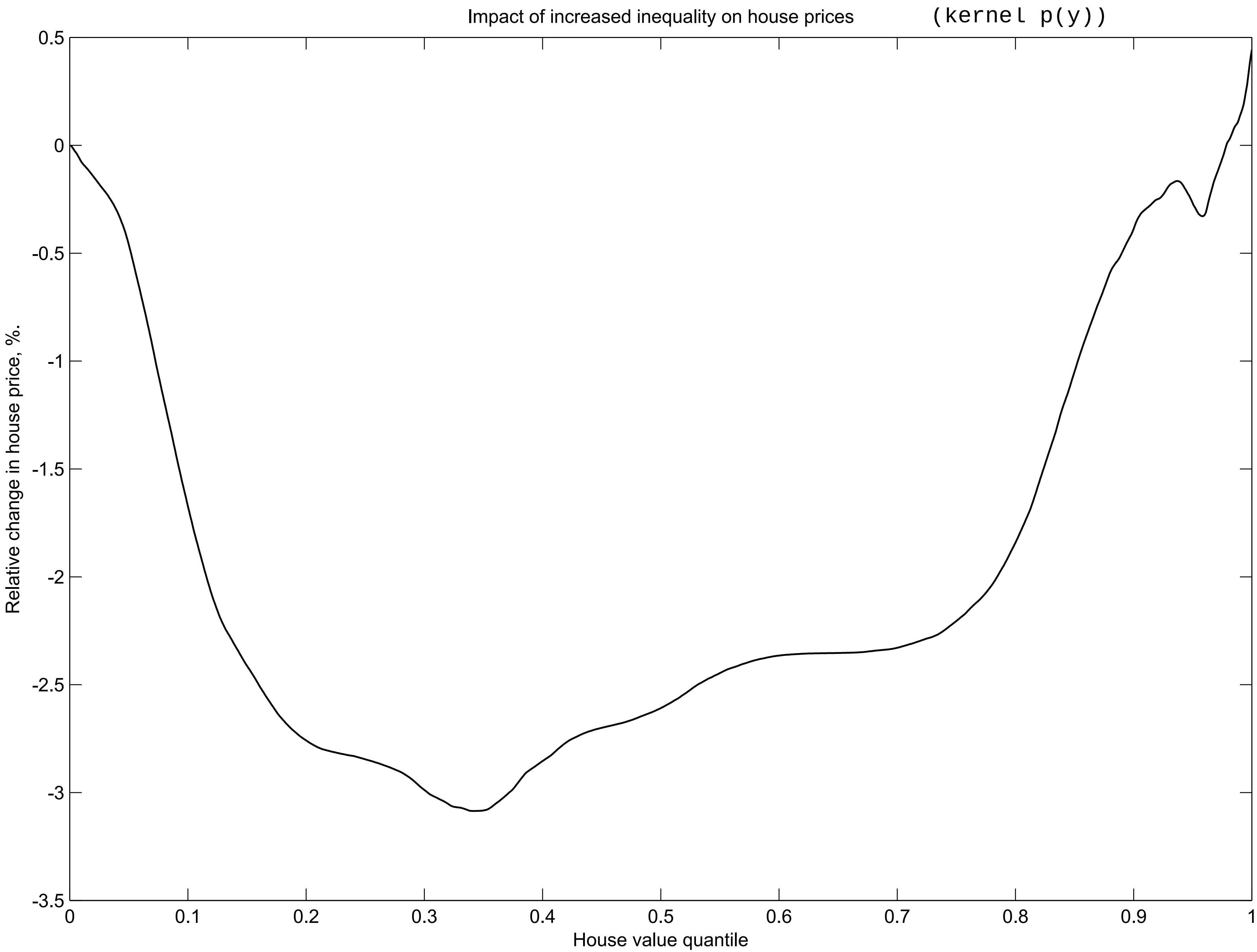

Impact of increased inequality on house prices (kernel p(y))
Relative change in house price, %.
House value quantile
0.5
0
-0.5
-1
-1.5
-2
-2.5
-3
-3.5
0
0.1
0.2
0.3
0.4
0.5
0.6
0.7
0.8
0.9
1

Figure 10

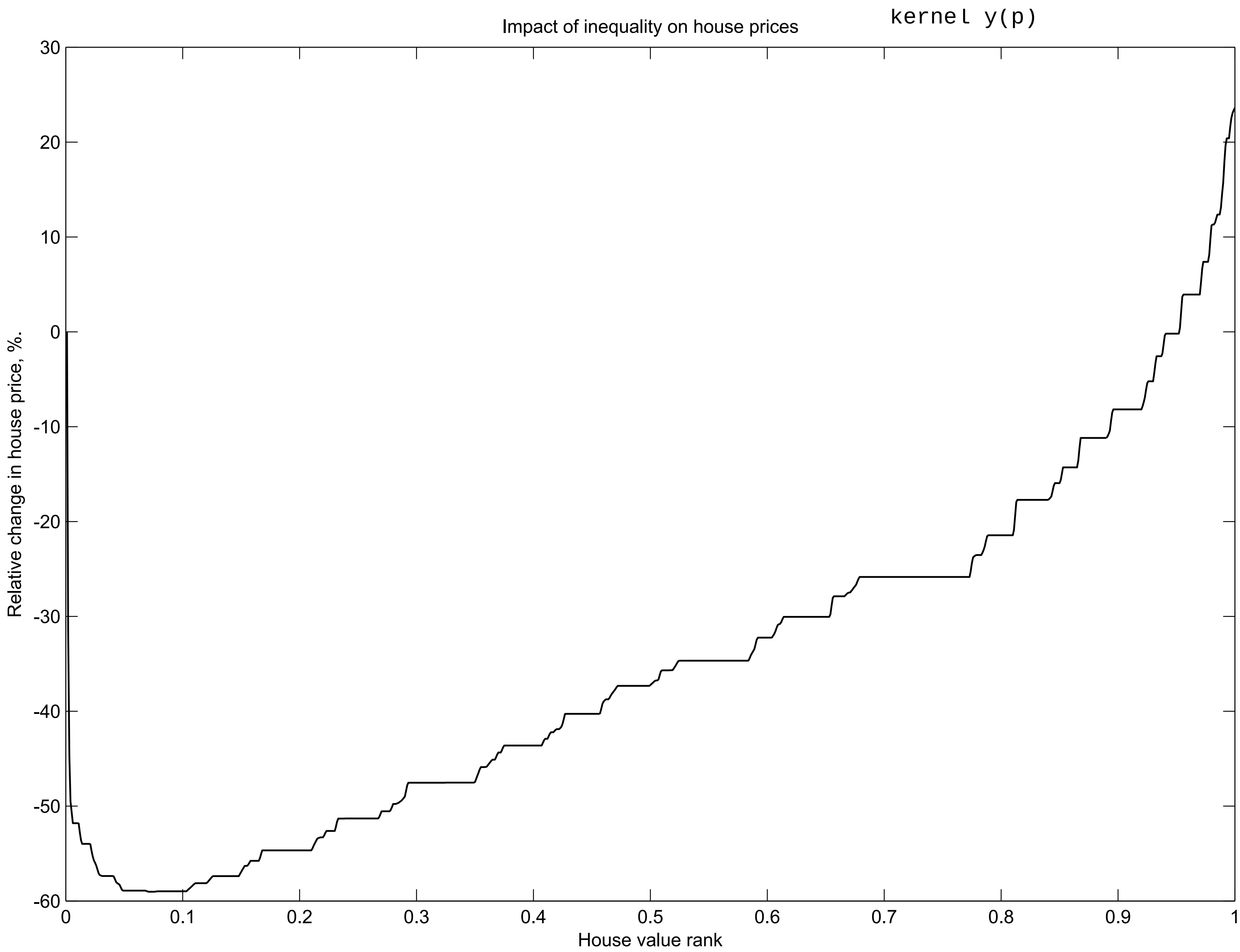

kernel y(p)
Impact of inequality on house prices
Relative change in house price, %.
30
20
10
0
-10
-20
-30
-40
-50
-60
0
0.1
0.2
0.3
0.4
0.5
0.6
0.7
0.8
0.9
1
House value rank

Figure 11

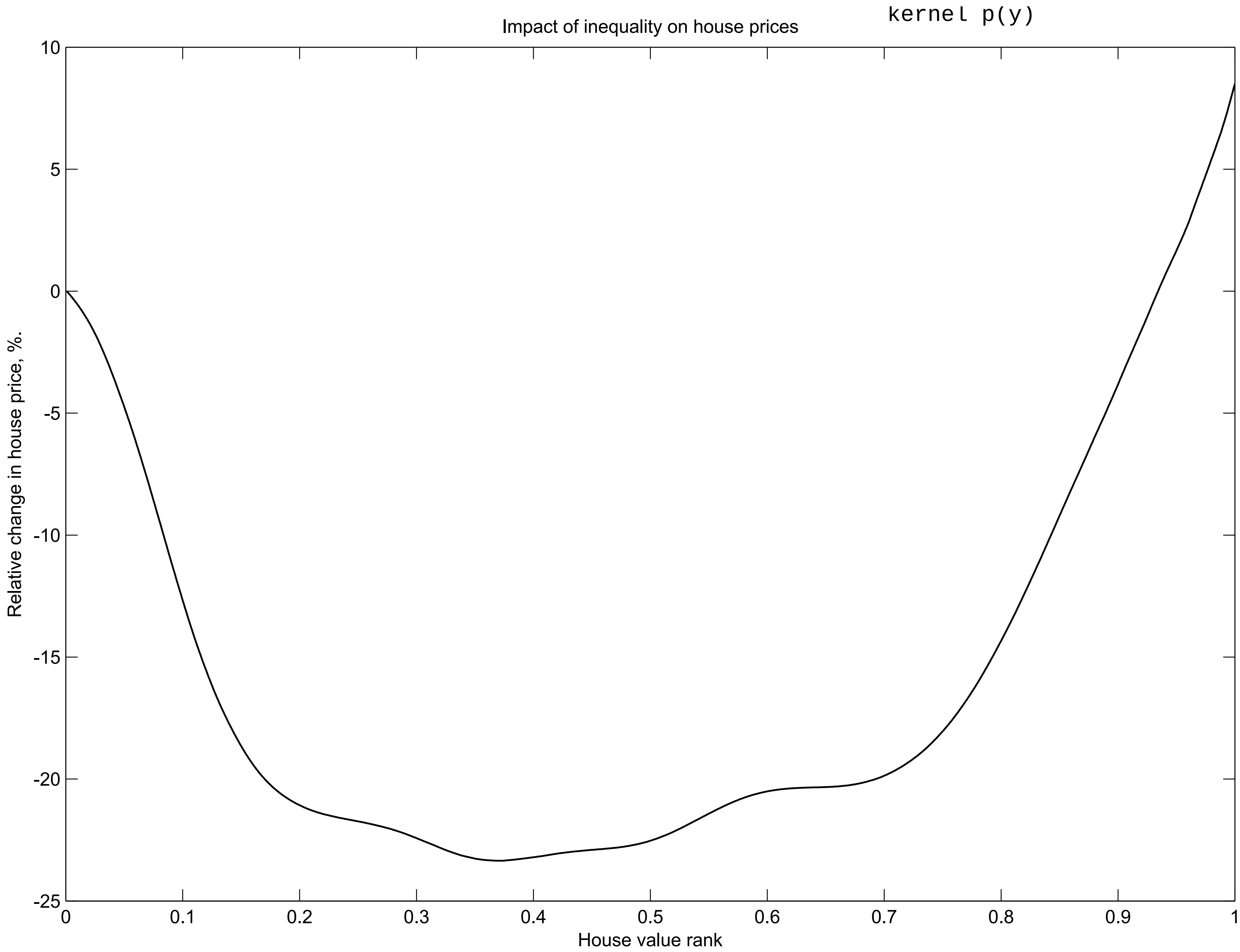

kernel p(y)
Impact of inequality on house prices
Relative change in house price, %.
10
5
0
-5
-10
-15
-20
-25
0
0.1
0.2
0.3
0.4
0.5
0.6
0.7
0.8
0.9
1
House value rank

Figure 12