C | E | D | L | A | S

Centro de Estudios
Distributivos, Laborales y Sociales

Maestría en Economía
Universidad Nacional de La Plata

Implicit Rents from Own-Housing and Income Distribution: Econometric Estimates for Greater Buenos Aires

Leonardo Gasparini y Walter Sosa Escudero

Documento de Trabajo Nro. 14
Octubre, 2004

Implicit Rents from Own-Housing and Income Distribution

Econometric Estimates for Greater Buenos Aires

Leonardo Gasparini[*]
Universidad Nacional de La Plata, Argentina

and

Walter Sosa Escudero
Universidad de San Andrés and
Universidad Nacional de La Plata, Argentina

This version: April 14, 2004

Abstract

Most income studies do not take into account the implicit rent obtained by households who inhabit their own dwellings, a fact that introduces a potentially relevant bias in inequality, poverty, and welfare measures. In this paper we estimate these implicit rents for the Greater Buenos Aires area from information of Argentina's National Household Expenditures Survey (ENGH) of 1996/7. Based on a sample of households that rent their dwellings, quantile regressions are used to estimate observed rents from a hedonic model. Estimated coefficients are applied to households that do not rent their houses or apartments in order to predict the implicit rent derived from living in an owned house. Estimated implicit rents are added to the standard notion of household income and various inequality measures are re-estimated. We find that the consideration of these implicit rents reduces inequality due to an income elasticity of spending in housing less than one, and to the relatively large proportion of house owners in the lower strata of the income distribution.

---

[*] Corresponding author: Leonardo Gasparini, Centro de Estudios Distributivos, Laborales y Sociales (CEDLAS), Department of Economics, Universidad Nacional de La Plata, calle 6 entre 47 y 48, oficina 516, (1900) La Plata, Argentina. Phone/fax: 54-221-4229383. E-mails: lgmm@netverk.com.ar and wsosa@udesa.edu.ar. A previous version of this paper was written in the framework of the Project of Technical Assistance for the National System of Public Investment, Dirección Nacional de Programación del Gasto Social, Ministerio de Economía y Obras y Servicios Públicos, Argentina. We thank the efficient assistance of Verónica Alaimo, the comments of Marcela Harriague, Sebastián Galiani, Omar Arias and two anonymous referees, and the bibliographical suggestions of Judith Yates. Georgina Pizzolito helped with the translation from an earlier version in Spanish. Errors and omissions are of our responsibility.

1. Introduction

The concept of income refers to the flow of resources obtained as remuneration to the use of all the assets owned by an individual or household. According to this definition, income should include not only the returns for the use of labor and capital, but also any other rents produced by the possession of durable goods, such as houses or cars. Families living in their own dwellings implicitly receive a flow of income equivalent to the market value of the service that the use of this property represents for them. This remuneration should be computed as part of household income, even though it is never recorded in a formal market and it is not usually registered as income in household surveys.

The main goal of this paper is to estimate the implicit rent received by those families that inhabit their own dwellings, and then study the distributional impact of adding these rents to the standard notion of income as it is usually reported to household surveys. The paper presents a discussion of the alternative estimation procedures and the corresponding empirical difficulties found in the implementation of these methods. The analysis is based on information for Greater Buenos Aires (GBA) of Argentina's National Household Expenditures Survey (ENGH) for 1996/7. This data set includes information on individual income and some characteristics of the dwellings, like number of rooms, bathrooms, construction material, availability of water connections, heating system, etc. Additionally, we observe the value paid by families that rent their houses and apartments.

The logic used to estimate the implicit rents is simple. Using the sub-sample of families that pay a rent, we estimate a hedonic model of rents as a function of observable characteristic of the housing. Then, we predict how much a family living in a household of their own would pay, conceptually, in terms of an "implicit rent" on the basis of observable characteristics. These estimations of the implicit rent are then added to the standard reported notion of household income.

Standard approaches based on mean regression models can be appropriate to estimate the average rent for the whole population. Instead, if as it is the case of this paper, the goal is to study the distributional effects, standard regression methods can lead to inaccurate estimations. This paper uses quantile-regression methods where a hedonic equation is estimated for each quantile of the conditional distribution of rents.

The analysis is completed with a comparison between the original income distribution and the resulting one corrected by imputing these implicit rents. In the following section we discuss with more detail the problems caused by ignoring the rents derived from living in an owned house or apartment, and the methodology we follow to

approach this problem. In Section 3 we discuss the estimation and inference strategies. Section 4 describes the sample, presents the results, and evaluates the quality of the estimated models. An assessment of the impact of including implicit rents on the income distribution is presented in Section 5. Section 6 concludes with some final comments.

2. The problem

In this section we present a very simple model to highlight the importance of adjusting the standard notion of income by the implicit rents of living in an owned dwelling. Let us assume an economy composed by individuals indexed with i. At the beginning of their active life these agents inherit H_i and decide whether to purchase a house. The budget constraint is

(1) $$H_i - c_i(1-\alpha_i)p(q_i) = s_{i0}$$

where c_i is a binary variable that takes value 1 if the individual i buys the house and 0 otherwise, $\alpha_i \in [0,1]$ is the proportion of the house price financed with a loan, q_i is a vector of characteristics of the house, $p(q_i)$ is a function that indicates the market price of the house, and s_{i0} is the amount saved or borrow by the individual.

During her active life the individual earns income Y_i, consumes x_i, pays a rent $A(q_i)$ in case of not having bought a house in the previous period, and repays the loan in case of having bought the house with a loan. Assuming a unique market interest rate r, the relevant budget constraint is:

(2) $$Y_i + s_{i0}(1+r) = x_i + (1-c_i)A(q_i) + c_i\alpha_i p(q_i)(1+r) + s_{i1}$$

Finally, person i retires and leaves her savings (or debts) and what is obtained from selling the house as inheritance to the next generation (R_i). With a stable price function and no depreciations, the new budget constraint is written as,

(3) $$s_{i1} + c_i p(q_i) = R_i$$

Combining (1), (2) and (3) the following budget constraint can be obtained:

(4)

$$Y_i + \left[H_i - c_i(1-\alpha_i)p(q_i)\right](1+r) + c_i p(q_i) = x_i + (1-c_i)A(q_i) + c_i\alpha_i p(q_i)(1+r) + R_i$$

If the individual chooses to buy the house with her own resources (c_i=1, α_i=0), the relevant budget constraint becomes

(5)
$$Y_i + H_i(1+r) - rp(q_i) = x_i + R_i$$

If she decides to buy the house with a loan (c_i=1, α_i>0), under our assumptions the budget constraint turns out to be equal to (5), while if she rents the house (c_i=0), it becomes

(6)
$$Y_i + H_i(1+r) = x_i + A(q_i) + R_i$$

If the housing market is competitive, prices will adjust for equation (7) to hold.

(7)
$$p(q_i) = \frac{A(q_i)}{r}$$

This equation implies indifference for similar individuals between buying or renting a house, since if (7) holds, equations (5) and (6) become equivalent. Living standards of similar individuals are identical, regardless of the decision of buying or renting a house, and regardless of the decision of buying it with own funds or with a loan.

In practice, studies based on household surveys estimate an individual's living standard by her household current income or expenditures adjusted by family structure. In this simplified example, the survey would register as current income $Y_i + r.(H_i - p(q_i))$ for those owners who bought the house with own resources, and $Y_i + rH_i$ for tenants. Clearly, this standard practice underestimates the income of house owners, introducing a bias in the measurement of welfare and inequality.[1]

An alternative to alleviate this problem is adding to the owner's income an estimate of the interest payments lost as a consequence of having bought the house. From equation (7) this value is equal to the savings that arise from not having rented a house similar to the one she owns.

If the house was bought with a loan, some possibilities arise. If there is a unique interest rate r, and in the surveys people report $Y_i - \alpha_i.r.p(q_i) + r.(H_i - (1-\alpha_i)\ p(q_i)) = Y_i +$

---

[1] A similar phenomenon occurs when approximating living standard using expenditures, since surveys

$r.(H_i - p(q_i))$ as current income, then adjusting income for the implicit rent of own-housing alleviates the problem. However, some people may report gross income, ignoring the term $\alpha_i.r.p(q_i)$. In that case, adding the implicit rent would overestimate income. Even when people report net income, if interest rates for loans and deposits differ, adding estimates of the implicit rent from own-housing would not completely solve the problem of comparability among incomes. Finally, if in contrast to the simplified model presented above the individual repays the loan in several periods, the fact that she has pending payments implies under most institutional settings that she is the owner of only a fraction of the dwelling.

The approach outlined in this section relies on the assumption that the rental housing market is competitive. Although perfect competition is rarely attainable, in many cases it seems to be a reasonable approximation to housing markets. For instance, in Greater Buenos Aires the housing market is characterized by multiple suppliers and low government intervention. Social housing is not widespread in Argentina. A Living Standard Measurement Survey conducted in 1997 (Encuesta de Desarrollo Social) reveals that only 0.5% of the households in the Greater Buenos Aires live in subsidized social housing, while 5% obtain credit from a public institution to build a house. Presumably, only a small fraction of this 5% gets loans at subsidized interest rates. Additionally, although rent controls were common in other decades, at the beginning of the 1990s the housing market was deregulated, and subsidies to tenants and rent controls eliminated.

3. Estimation strategy

We have argued for the need to modify the owners' incomes by adding the implicit rent from their dwellings. A basic problem is that these rents are not directly observable and, hence, must be estimated. The choice of an appropriate method eventually depends on the level of aggregation desired and on data availability. There are several antecedents in the literature that are relevant for this work. Katz (1983) reviews alternative methods to value consumer durable goods and their rents. The problem of assessing the aggregate value of implicit rents of owner-occupied dwellings can be approached from a 'national accounts' perspective, where aggregate rental values of ownership are computed based on national accounts data. Yates (1994) is a relevant reference since it discusses several strategies to compute rents based on the national accounts approach. In order to study distributional issues it is necessary to impute rents at the individual

usually report x_i as current expenditures for owners and $x_i+A(q_i)$ for tenants.

level. To this purpose, Yates (1994) proposes to apply average rates estimated from the national accounts to individual level of dwelling value. These rents may obey either a market value criterion, where rates are computed based on observed market rental rates, or an opportunity cost one, where rates are computed based on the return to alternative investments to owning a house or apartment.

Alternatively, as hinted by Malpezzi (2000), we will follow a micro-data approach based on the fact that the available survey provides two relevant pieces of information: housing characteristics for every household in the sample, and the rents paid by tenants. This information allows us to estimate equations that relate housing characteristics to the rents paid, and then to impute the estimated parameters to the housing characteristics inhabited by owners.[2]

Technically, the estimation exercise corresponds to the family of hedonic price models: housing is seen as a compound good that results to be the aggregate of a group of characteristics (number of rooms, location, availability of garden, etc.), and its price is the aggregate value of these characteristics (Rosen, 1974). The estimation of hedonic functions is likely to be one of the most thoroughly researched topics in applied econometrics, since it faces the analyst with a variety of textbook difficulties including model selection, choice of functional form, potential heteroscedasticity, measurement errors, etc. A complete review of these topics far exceeds the goals of this paper. The study of the hedonic price functions dates back to the pioneer work of Waugh (1929), who characterizes the quality of a product as function of its attributes and estimates the implicit price of each characteristic. However, the term "hedonic" has not been used until the work of Court (1939) and was popularized among economists by Griliches (1961). The eighties and nineties have witnessed an enormous amount of empirical work that estimates hedonic models with two main purposes: to infer the implicit prices of the attributes of a house, and to predict its value. Estimations have followed both parametric (see Cropper, Deck and McConell (1988), Craig, Kohlase and Pappel (1991), and Laakso (1997)) and non-parametric approaches (Pace, 1995). The survey by Sheppard (1999) focuses on different econometric problems that face hedonic estimations and some directions to overcome them. Restrictions are often related to data availability. For example, although the theoretical literature on urban economics

---

[2] The methodology requires all houses to be either rented or owned by its occupiers. However, it could be the case that some non-owners do not pay rents. If this behavior is widespread and not uniformly distributed in the population (e.g. more usual for the poor), then a bias in the estimations could arise. Although we do not deny this possibility, in practice for the Greater Buenos Aires it does not seem to be very relevant. From the survey EDS 1997 the fraction of people who are either renters or owners is more than 97%. The rest comprises from poor people who occupy lots in slums to rich managers or government officials who get expensive houses from their jobs while in office.

suggests including location as an explanatory variable, in many empirical papers this variable is not available.

A conclusion that emerges from this literature is that the choice of an appropriate econometric strategy strongly depends on the intended purpose of the analysis. As stressed by Sheppard (1999, pp. 1614) hedonic models are used with two main purposes. The most popular and more challenging one is to recover the implicit (hedonic) prices of attributes, and the second one is to forecast the value of properties. It is crucial to remark that for the purposes of this paper it is the second use of hedonic models that guides the choice of a relevant framework. In general terms such models estimate a simple function of the log of prices (rents paid) as a function of a variety of explanatory variables related to different observable housing characteristics.

The standard approach estimates mean regression models, which have serious limitations to study distributional effects. In particular, as it will be shown in the empirical section, the standard method predicts relatively high rents for poor households and relatively low rents for the rich ones. As an example, consider the case of poor households. Given that most of the explanatory variables of the hedonic model are binary variables indicating whether the house or apartment has a certain characteristic (electric doorbell, swimming pool, garage, etc.) or variables with larger values for "better" characteristics (number of rooms, bathrooms, etc.), for the very low income households these indicators are zero or close to zero, for which the prediction of the rent is, essentially, the intercept of the hedonic function. In the standard interpretation, the intercept of a linear model is the average of the market value of these non-observable characteristics of the housing. In fact, if none of the characteristics were observed, the intercept will be the average rent paid by all the individuals in the sample. Then, the prediction of a standard regression model would assign to all households the same average value of these non-observable characteristics, spuriously inflating the value of the housing inhabited by the poorest households, and underestimating the value of the housing for high-income families.

In order to overcome this problem we use quantile regression methods, where a hedonic equation is estimated for each quantile of the conditional income distribution. Koenker and Hallock (2001) is a recent and elementary introduction to these methods. The implicit assumption is that the monetary value of the demand for non-observable characteristics is related monotonically to the distribution of income, so that conditional quantiles of the distribution of non-observables coincides with the quantiles of the distribution of income, that is, individuals in the first decile of the distribution of income are assigned as their demand for non-observables the first decile of the conditional distribution of non-observables, and so on. Though restrictive, this

assumption help us distribute non-observables in a less drastic way, since poorer households will be assigned a relatively lower value of the non-observables.

Formally, the model to be estimated has the following form:

$$Q(A/\tau) = q\beta_\tau + Q(u/\tau)$$

where $Q(A/\tau)$ is the τ-th quantile of the conditional distribution of rents *(A)*, q is a vector of observable housing characteristics, including a constant, β_τ is a vector of coefficients that indicates how observable characteristics affect the τ-th quantile of the conditional distribution of *A,* and u is an error term that contains the market values of all non-observable characteristics of the paid rents. It is important to note that such specification permits that explanatory variables affect in different ways the conditional quantile of the explained variable. In particular, each quantile has its own intercept. In the basic regression quantile model initially proposed by Koenker and Bassett (1978) the coefficients can be consistently estimated as a solution to the following linear program:

$$\hat{\beta}_\tau = \arg\min_\beta \sum_{i=1}^{n} \rho_\tau (A_i - q_i\beta)$$

with:

$$\rho_\tau(v) = v(\tau - I(v < 0))$$

where $I(\ )$ is an indicator function. With the purpose of testing hypothesis and constructing confidence intervals, let $\hat{\beta}_n \equiv (\hat{\beta}_{\tau_1}, \ldots, \hat{\beta}_{\tau_m})$ be a *km* vector of *k* coefficients estimated for *m* different quantiles based on a sample of *n* observations, and let β_o be its population counterpart. Under some regularity conditions (Koenker and Bassett, 1978) it can be verified that:

$$\sqrt{n}(\hat{\beta}_n - \beta_0) \Rightarrow N(0, \Omega \otimes Q_0^{-1})$$

where Ω is a *m* x *m* matrix with typical element:

$$\omega_{ij} = \frac{min(\tau_i, \tau_j) - \tau_i\tau_j}{f(F^{-1}(\tau_i))f(F^{-1}(\tau_j))}$$

$Q_0 = \lim_{n \to \infty} n^{-1}(Q'Q)$, "$\otimes$" is the Kronecker product, and "$\Rightarrow$" denotes convergence in distribution. For practical purposes it is necessary to estimate the parameters $f(F^{-1}(\tau))$

using standard non-parametric density estimation techniques. This result can be easily used to build confidence intervals.

Model building in hedonic analysis is a complicated issue. A first problem is the choice of a relevant set of explanatory variables used to predict rents. As it is the case of compound commodities which are the subject of hedonic analysis, rents are treated as a "price" paid for the use of all the characteristics (observed or not) embodied in the dwelling. Consequently, theory places few restrictions regarding which variables to include or exclude from the analysis since all observed characteristics are implicitly valued by consumers. From an empirical point of view, as it is well known, the price to be paid for incorrect variable exclusion is in terms of potential biases, and the benefits are related to reductions in the dimension of the parameter space which might eventually translate into efficiency gains. The strategy of avoiding possibly non-realistic theory restrictions points toward including as many variables as they are available, which for a given sample size, may induce high multicollinearity among the regressors. This might seriously affect the precision of estimated hedonic prices due to the well established fact that high multicollinearity affects the identification of the true hedonic coefficients. But if, as it is the case of this paper, the goal is to predict prices from a given set of explanatory variables, high multicollinearity is also well known to not affect the forecast performance of standard OLS estimation (Judge, *et al.*, 1995, pp. 897 and 901). Hence, for the purposes of this paper high multicollinearity or proper variable selection is not a serious problem.

Alternatively, the quantile methods used in this paper require to keep the dimension of the problem at a relatively low level. An arbitrary high number of regressors seriously affect the computational performance of the algorithms used to produce estimations, besides affecting the precision of estimates as in the case of OLS methods. Hence, in order to keep the dimension of the problem as low as possible without harming the predictive performance of the model, the estimation strategy used is the following. In a first exploratory stage a simple regression model is estimated for rented houses and apartments, incorporating all the explanatory variables. Considering that these variables are proxies for economic factors not directly observable, and with the aim of obtaining a simplified representation, a stepwise search method was used. The search was implemented "forwards" (beginning with a model without explanatory variables, adding them according to its statistical significance measured by their 't'-statistics) and "backwards" (beginning with all the explanatory variables and eliminating them one-by-one according to the same criterion). The use of a search algorithm is arbitrary, but no less than other commonly used strategies like plain variable deletion or the use of principal component regression (Judge *et al.*, pp. 909).

For the predictive goals of this paper, the nesting structure of the original and final model implies that the ultimate validity check is that the low-dimensional model adopted does not imply a significant loss in predictive power compared to the high-dimensional original model. To this purpose we compare standard goodness-of-fit criteria to assess the loss in explanatory power between the original model and the chosen one.

In a second stage the simplified versions of the models are estimated using quantile regression methods. Surely this strategy is subject to some caveats. In particular, there is no guarantee that the model chosen by OLS methods would coincide with the one chosen by a similar strategy implemented base on quantile regressions. The main reason to proceed to use OLS in the exploratory stage is, precisely, to reduce the dimension of the problem so as to avoid the computational disadvantages of a high dimensional parameter space in quantile regression, which would make it impossible to carry out the search procedure. Fortunately, as will be discussed in the next section, the slope coefficients estimated by OLS and regression quantiles at different quantiles do not differ significantly, which suggests that both methods agree with the selected model.

The choice of the log functional form is based mostly to be in accordance with a substantial part of the previous literature. This is a relevant topic that might affect the estimation of hedonic prices (but not necessarily the prediction of aggregate prices), as documented by Cropper, Deck and McConell (1988). A full assessment of these biases in the quantile regression framework is an interesting topic for further research.

4. Estimation of hedonic rent models

The sample

We use data from the Argentina's National Household Expenditures Survey (ENGH) conducted between February 1996 and March 1997.[3] We work with a sub-sample corresponding to the metropolitan area of Greater Buenos Aires (GBA).[4] The survey includes information on household expenditures and incomes, and on different demographic, occupational and educational characteristics of the population. An important subset of variables is related to dwelling characteristics. In the estimation stage we use a sub-sample of households that report living in rented houses or

---

[3] The ENGH is the only survey in Argentina with information on household incomes, rents and housing characteristics. The other available surveys (EPH, EDS and ECV) do not report rents paid by tenants.

[4] The Greater Buenos Aires comprises the city of Buenos Aires and its suburbs (around 2,600 km^2). It is an area inhabited by 12 million people, living in around 3 million residences.

apartments for which we observe a positive rent. This sub-sample includes 544 households -201 houses and 343 apartments- out of the 4112 households in the whole survey for GBA.

We estimate different models for houses and apartments, which allow explanatory variables to have differential effects according to the type of housing. Table 1 presents means and standard deviations of the variables used in the analysis.

[INSERT TABLE 1 HERE]

Three variables usually identified as potential determinants of the price of a house are not reported in the ENGH: size, age and location. The first variable can be approximated by the number of rooms and bathrooms, and the presence of a garage, kitchen and garden. However, it is difficult to estimate the age of the house or its location with the variables shown in Table 1.[5]

OLS estimations: initial and simplified models

Having divided the sample in houses and apartments, we proceeded with the estimation of a simple model of rents paid as a function of observable characteristics of the dwelling. The variable of interest *(rent)* takes values between 50 and 5000 pesos per month. As it is standard we have worked with the natural logarithm of rents *(lrent)* as the dependent variable.

[INSERT TABLE 2 HERE]

OLS estimations of the initial and simplified hedonic model for houses are shown in Table 2. The first two columns present results for houses. The *F*-tests suggest that the explanatory variables are significant to explain the logarithm of rents. In the initial model the R^2 indicates that the linear model explains approximately 50% of the variability of the log of paid rents, which is reasonably high for micro data. With a relatively conservative level of significance (10%), both search methods (backwards and forwards) select exactly the same model based on the same sub-group of explanatory variables.

These results suggest that a simple model of rents can be based on seven observable characteristics: number of rooms, telephone, running water, security, parking, hot water in the kitchen and electric doorbell. This simplified model is

---

[5] The availability of a garden may be a proxy for distance from the city center.

significantly more parsimonious than the original one that includes 22 explanatory variables. Formally, an *F*-test of linear restrictions indicates accepting the null hypothesis implicit in the simplification. The loss of explanatory power is, consequently, of very little significance. The R^2 of the simplified model is 0.47, while that of the original one is 0.50. This effect is better captured by the adjusted R^2. In fact, the adjusted R^2 of the simplified model (0.45) is slightly higher than the one corresponding to the original one (0.44).

Although for this paper the main objective of the estimation is predicting rents, it is useful to discuss some particular results of the simplified model. All the estimated coefficients have the expected signs. An additional bedroom increases the value of the rent 17%. Surprisingly, other size variables turn out to be non-significant, such as the number of bathrooms. Telephone, water, hot water in the kitchen, and electric doorbell can be interpreted as indicators of the quality of the housing. It is interesting to observe that these characteristics increase the value of the rent in approximately the same magnitude. This fact suggests that, at least in terms of the determination of the value of the rent, an elementary index of quality of the housing can be built as the aggregation of these characteristics. Consequently, according to the obtained results, the marginal contribution of any of the mentioned characteristics in the value of the rent is always the same. The value of the coefficient of the private security variable is, although not precisely estimated, significantly high. Taking into account the low proportion of observations with this indicator equal to 1, it is possible that this variable identifies dwellings of very high quality. Finally, the availability of a garage increases the value of the rent 14%. This variable has a less clear interpretation, since, as we have previously remarked, a garage can be interpreted as an indicator of both the size and the quality of the house.

The last two columns of Table 2 present similar results for apartments. As in the previous case, the *F*-test suggests that explanatory variables are, as a group, significant to explain rents. The R^2 coefficient indicates that the linear model explains approximately 40% of the variability in the dependent variable. The problem of high correlation between explanatory variables seems to be relevant in this case, and complicates the identification of the parameters and, consequently, the selection of a relevant simplification. Unlike the case of houses, different stepwise procedures lead to different specifications. Keeping in mind that the presence of near multicolineality does not affect the predictive ability of the model, and the fact that the R^2 is slightly higher than in the case of the downwards specification, the backwards model was selected as an acceptable simplification. Again, the loss in explanatory power is not significant with respect to the complete model. This indicates that, compensated by the gain in degrees-

of-freedom, the explanatory power of both models is virtually indistinguishable. Besides, the correlation structure of the variables maintained in the simplified model does not show values greater than 0.37, which suggests that the problem of near multicolineality in the original estimation does not exist in the simplified version.

Regarding the interpretation of the coefficients, the one corresponding to the number of bedrooms is very similar to the one obtained in the model for houses. From this fact it could be concluded that the contribution of an additional bedroom increases the value of the rent in approximately 17%, independently of whether it is a house or an apartment. The rest of the variables have the expected signs. Factors that significantly and positively contribute to the rent are the number of bathrooms, a heating system, and having access to the natural gas and sewerage network. Apartments located in a complex have lower expected rent values, while apartments in higher floors are on average more expensive.

Estimations based on quantile regressions

In terms of explanatory and predictive power, the R^2 of the estimated models indicate that the variables incorporated as regressors explain 40 to 50% of the total variability of the (log) rents. This suggests that the mechanism through which rents are determined has idiosyncratic, non-observable components that cannot be properly addressed by the large number of explanatory variables incorporated in the model. As it has been advanced in Section 2, the intercept of the linear regression model estimates the average of the market valuation of these unobserved heterogeneities, which seem to have a non-trivial role in the determination of rents.

Table 3 presents estimates of the quantile regression model for houses and apartments for quantiles 0.1, 0.25, 0.5, 0.75 and 0.9 of the conditional distribution of rents. Figure 1 summarizes the estimated coefficients. Each graph presents estimations of the coefficients of a variable for different quantiles. The dotted line corresponds to point estimates, and the solid lines to the 95% confidence intervals, using the methods described in the previous section.

[INSERT TABLE 3 HERE]

[INSERT FIGURE 1 HERE]

5. Imputed rent and income distribution

After estimating rent models for houses and apartments, we computed the predicted implicit rent values associated to housing ownership. In this section we discuss the results of these imputations and their impact on the measurement of income inequality.

Table 4 groups families by deciles according to their household per capita income prior to the adjustment by implicit rents. Column (i) reports the percentage of owners in each decile. Home ownership is very widespread in Greater Buenos Aires: more than 85% of the households inhabit in their own houses or apartments. It is interesting to note that owning a house or apartment is more common among low-income people, which suggests that, *ceteris paribus*, adjusting incomes by implicit rents would generate a fall in inequality measures.

[INSERT TABLE 4 HERE]

Columns (ii) to (vi) refer to the group of owners. Column (ii) presents the average total household income for each decile, while column (iii) shows the average imputed rents estimated by OLS. The average rent for the lowest income decile is $241 while for the top decile this value is $542. The estimated imputed rent as a proportion of total family income is strictly decreasing in income (62% for the poorest decile and 16% for richest one). This result is consistent with an income elasticity of housing expenses less than unity, which, again, suggests that the adjustment of incomes by implicit rents would induce an equalizing effect on the income distribution. This result is in line with what previous research has found in local markets (see Malpezzi, 2000 and Malpezzi and Mayo, 1987).

The results in columns (iii) and (v) must be cautiously interpreted, keeping in mind the following considerations. First, the estimated value of the implicit rent as a proportion of income might seem to be too high. However, it should be taken into account that in most surveys incomes are more under reported than some expenditure items, like housing rents, for which the figures in column (v) (based on incomes reported to the survey) are probably higher than the actual ones. Gasparini (1998) estimates a level of under report in the ENGH which on average implies multiplying reported incomes by a factor of 2.5. Naturally, this adjustment would imply a large fall in the share of implicit rents with regard to the results of column (v) in Table 4.

Second, the model estimated to predict rents is based on limited information, since it is not possible to obtain data on some important rents determinants, like the location of the house or the exact size of the dwelling. The estimated model tends to

capture the influence of these characteristics in an indirect way through variables related to these non-observable factors in the sample. This implies a greater variance since the error term incorporates factors that are not directly observed which, as it has been discussed in the previous section, reduces the predictive ability of the model. This suggests that, given the limitations of the sample information, it would be cautious to assign these estimations a certain rank of variability caused by the impossibility of measuring all the relevant effects in the rent determination. However, it is important to note that the goodness-of-fit measures obtained in the estimation are not significantly lower than the ones obtained in similar studies for the local market (Figueroa and Lever (1992), Stumpf Gonzalez and Torres Formoso (1997) and Gomez Mera (1998)), even though these studies are based on a larger set of rent determinants obtained directly from the real estate market.

Third, it is interesting to consider to what extent housing characteristics differ between rented and non-rented dwellings. In Table 5 we present a summary of the main housing characteristics for owner and tenant households for both apartments and houses, and a formal test for the null of no differences in their means.[6] In general, even when owner occupied houses and apartments are on average larger than rented ones, there is considerable overlap between the distributions of the characteristics of both samples.

[INSERT TABLE 5 HERE]

Additionally, it is interesting to compare the predicted imputed rents with the rents paid. While the average imputed rent is $379, the average rent reported in the survey is $365. According to columns (iii) and (viii) of Table 4, differences by deciles are not substantial.

These considerations suggest that the seemingly high values of the predicted implicit rent are mostly due to the income under-report phenomenon previously described, and not to the structural deficiencies of the model used for prediction or to differences in the characteristics between the households used for the estimation and those used for the prediction.

Table 6 presents intercepts of the conditional distribution of rents by per capita household income deciles. As mentioned in Section 3, instead of assigning a single intercept for every household (the one estimated by OLS), we use for each household the intercept corresponding to its relative position in the conditional distribution of

---

[6] Standard tests for differences in means were computed for count variables (rooms, bathrooms, etc.) and differences in proportions for binary indicators.

rents, under the assumption that the deciles of both distributions can be matched, that is, for households in the first decile we assign an intercept corresponding to the first decile in the conditional distribution of rents, and so on. For houses, the intercept is $101.6 for decile 1, and $240 for decile 10. This should be compared with the OLS intercept, $136.4. This implies that the adjustment in the intercept is more important for high-income families. In the case of apartments exactly the opposite occurs. The predicted intercept for the lower decile is $30.8 while the initial method predicted $77.3. This implies a correction of approximately 60%.

[INSERT TABLE 6 HERE]

The implicit rents estimated with this correction are presented in column (iv) of Table 4. As it was expected, the correction of the intercept reduces the estimation of the rents for the lower deciles (from 1 to 4) and increases it in the top deciles. For example, the original method predicted a rent that resulted to be 62% of the total family income computed in the survey while the corrected method predicts 53%.

Adjustment by implicit rents and income inequality

The last goal of this paper is to assess the impact on the Greater Buenos Aires' income distribution of adding the estimated implicit rents from owning a house or apartment.[7] Graph 2 presents non-parametric kernel-based estimations of the original income distribution, and the resulting one after adding implicit rents using quantile-based estimations. Table 7 presents various inequality indexes computed on nine different distributions. In the first column we show inequality measures for the per capita household income distribution computed with the original data from the ENGH. In the following four columns we add to that distribution the estimations of implicit rent. Columns (ii) and (iii) are based on OLS estimates while in the next two columns quantile regression estimations are shown.

[INSERT GRAPH 2 HERE]

As discussed in section 2 a problem arises if some people bought their houses with loans and have pending payments. From the ENGH the proportion of owners

---

[7] Inequality has been significantly increasing in Argentina (and in the Greater Buenos Aires) for the last three decades. Once a low-inequality economy compared to the rest of Latin America, Argentina is now reaching inequality levels close to the Latin American mean (Gasparini, 2003).

having pending payments is 7%.[8] We do not have information on the share of the loan already paid by each family. For this reason we take two extreme alternatives. In the even columns we ignore the fact that some families have bought their houses with credits. These families are treated as owners and are assigned rents derived from housing ownership. On the other hand, in the odd columns, rents are not imputed to families that report having pending payments for their houses.[9]

Columns (vi) to (ix) reproduce these exercises multiplying the per capita household income from the ENGH by a factor of 2.5, in order to adjust for under-reporting in the survey.

[INSERT TABLE 7 HERE]

The Gini coefficient of the original household per capita income is 0.445. As a result of adding the estimated implicit rents (using OLS) to the owners the Gini coefficient falls to 0.410. As it is expected, this drop is smaller if the quantile-based estimations are used, if we do not impute rents to credit holders and, in particular, if the adjustment for under-reporting is made. Taking into account these considerations, the Gini coefficient still falls but the change is smaller: from 0.445 to 0.432. The rest of the inequality indicators behave in a similar fashion.

The fall in the inequality indicators is driven by two facts. First, in Greater Buenos Aires home ownership is somewhat more widespread in low-income strata, which implies that an implicit rent imputation proportional to income would imply an equalizing effect on the income distribution. Second, the imputations turn out to be non-proportional to income, reflecting an income-elasticity of housing expenses smaller than one, which adds a second equalizing effect on the income distribution.

6. Final comments

This paper presents estimations of the implicit rents from own housing based on microdata from the National Household Expenditures Survey of Greater Buenos Aires. The paper discusses the different methodological and empirical difficulties related to the estimation process, which should necessarily be considered to evaluate empirical results. The estimates of implicit rents are added to household income and inequality indexes are computed over the resulting distribution. Recorded inequality decreases as

---

[8] The pattern by household income is not clear. The share of households with pending payments by quintiles is 6.4%, 7.6%, 6.3%, 5.6% and 8.5%.
[9] Naturally, a bias could arise if the relative size of the pending payment differs by income strata.

implicit rents are included due to widespread housing ownership along the income distribution, and an income-elasticity of housing expenses smaller than one

There exist several dimensions in which this investigation can be extended. Obviously, information availability, like housing location and other variables absent in the ENGH, should contribute to improve the quality of the estimations. Also, as mentioned in section 3, it is important to confront the results of this work to alternative estimation strategies. In particular, it would be relevant to aggregate the micro-level estimates obtained by this work and then compare results with rents estimated using the national accounts approach, as discussed in Yates (1994). Though it is methodologically complex to harmonize micro information with aggregate data, this is an important validation exercise that should be implemented.

References

Court, A. (1939). Hedonic price indexes with automotive examples. En *The Dynamics of Automobile Demand.* General Motors.

Craig, S., Kohlhase, J. and D. Papell (1991). Chaos theory and microeconomics. An application to model specification and hedonic estimation. *Review of Economics and Statistics* 73.

Cropper, M., Deck, J. and K. McConnell (1988). On the choice of functional form for hedonic price functions. *Review of Economics and Statistics* 70.

Figueroa, E. and Lever, G. (1992). Determinantes del precio de mercado de los terrenos en el área urbana de Santiago. *Cuadernos de Economía*, 86, pp. 99-113.

Gasparini, L. (1998). La incidencia tributaria del sistema impositivo en la Argentina. En *La Reforma Tributaria en la Argentina*, FIEL.

Gasparini, L. (2003). Different lives: inequality in Latin America and the Caribbean. In The World Bank, *Inequality in Latin America and the Caribbean: breaking with history?*

Gomez Mera, Laura (1998). El valor social de las plazas. Mimeo, Universidad de San Andres.

Griliches, Z. (1961). Hedonic price indexes for automobiles: an econometric analysis of quality change. *The Price Statistics of the Federal Government,* 73.

Katz, A. (1983). Valuing the services of consumer durables. *Review of Income and Wealth*, 29, 405-428.

Koenker, R. and Bassett, G. (1978). Regression quantiles, *Econometrica*, 46, 33-50.

Koenker, R. and Hallock, K. (2001). Quantile regression: an introduction. *Journal of Economic Perspectives*, 15: 143-156.

Instituto Nacional de Estadística y Censos (INDEC) (1998). *Encuesta Nacional de los Gastos de los Hogares 1996/97. Región Metropolitana del Gran Buenos Aires, Resultados Provisorios. Vol. 1.*

Judge, G., Griffiths, W., Carter Hill, R., Lütkephol, H. and Lee, T. (1985), *The Theory and Practice of Econometrics*, 2nd edition, Wiley, New York

Laakso, S. (1997). *Urban housing prices and the demand for housing characteristics*. ETLA, The Research Institute of the Finnish Economy.

Malpezzi, S. (2000). Housing. In Grosh, M. and Glewwe, P., eds., *Designing Household Survey Questionnaires for Developing Countries*, The World Bank, Washington.

Malpezzi, S., and S. Mayo (1987). The demand for housing in developing countries. *Economic Development and Cultural Change* 35, 687-721.

Pace, R. (1995). Parametric, semiparametric and nonparametric estimation of characteristic values within mass assessment and hedonic pricing models. *Journal of Real State Finance and Economics* 11.

Rosen, S. (1974). Hedonic prices and implicit markets: product differentiation in pure competition. *Journal of Political Economy*, 82, 34-55.

Rosen, H. (1995). *Public Finance*. Irwin.

Sheppard, L. (1999). Hedonic analysis of housing markets. In *Handbook of Regional and Urban Economics*, vol. 3.

Stumpf Gonzalez, M. and Torres Formoso, C. (1997), Estimación de modelos de precios hedonicos para alquileres residenciales. *Cuadernos de Economía*, 34, 101, 71-86.

Waugh, F. (1929). *Quality as a determinant of vegetable prices*. Columbia University Press.

Yates, J. (1994). Imputed rent and income distribution. *Review of Income and Wealth* 40, 43-66.

Table 1: Variables name and description

Description	Variable	Houses		Apartments	
		Mean	Std. Dev.	Mean	Std. Dev.
Rent	rent	373.04	459.47	371.95	0.36
Number of bathrooms	bathrooms	1.10	0.70	1.28	0.71
Number of bedrooms	bedrooms	2.75	0.90	2.29	0.91
Construction material (categorical)	material	2.62	0.72	2.95	0.26
Running water (1=yes, 0=no)	water	0.75	0.44	0.97	0.18
Hot water in the kitchen (1=yes, 0=no)	hot water	0.61	0.49	0.87	0.34
Air conditioner (1=yes, 0=no)	AC	0.05	0.22	0.08	0.28
Bathroom of exclusive use (1=yes, 0=no)	exc bath	0.99	0.12	0.99	0.12
Sewerage system (1=yes, 0=no)	sewerage	0.54	0.50	0.93	0.26
Parking (1=yes, 0=no)	parking	0.26	0.44	0.16	0.37
Kitchen (1=yes, 0=no)	kitchen	0.97	0.17	0.98	0.14
Stove (1=yes, 0=no)	stove	0.65	0.48	0.92	0.27
Belongs to a complex (1=yes, 0=no)	complex	0.01	0.11	0.05	0.22
Type of bathroom (categorical)	type bath	0.54	0.50	0.93	0.26
Distance from public transportation stop	transportation	0.93	0.25	1.00	0.07
Natural gas (1=yes, 0=no)	gas	0.84	0.37	0.96	0.19
Garden (1=yes, 0=no)	garden	0.52	0.50	0.23	0.42
Electric doorbell (1=yes, 0=no)	e.doorbell	0.08	0.27	0.73	0.45
Water source (categorical)	water source	0.71	0.45	0.97	0.18
Security (1=yes, 0=no)	security	0.02	0.13	0.08	0.28
Heating system (1=yes, 0=no)	heating	0.44	0.50	0.64	0.48
Telephone (1=yes, 0=no)	phone	0.41	0.49	0.72	0.45
Public lighting (1=yes, 0=no)	lighting	0.95	0.22		
Elevator (1=yes, 0=no)	elevator			0.58	0.50
Floor 2	floor 2			0.15	0.35
Floor 3	floor 3			0.12	0.33
Floor 4	floor 4			0.37	0.48
Floor 5 or more	floor 5			0.18	0.38
Doorkeeper (1=yes, 0=no)	doorkeeper			0.50	0.50

Source: Authors' calculations based on the ENGH.

Table 2: OLS Estimation

	Houses		Apartments	
	Initial model	Simplified model	Initial model	Simplified model
bedrooms	0.1672	0.1711	0.1793	0.1747
	(3.8390)	(4.1630)	(7.5760)	(7.5710)
lighting	-0.0258			
	-(0.2420)			
water	0.1641	0.2304	0.2380	
	(1.1260)	(3.5740)	(1.9710)	
sewerage	-0.1268		-0.1288	
	-(1.2470)		-(1.1210)	
gas	0.0546		-0.1996	
	(0.5930)		-(1.5000)	
complex	0.0121		-0.1466	-0.1434
	(0.0710)		-(1.6800)	-(1.7080)
e.doorbell	0.2788	0.2409	0.0058	
	(2.5770)	(2.2160)	(0.0950)	
security	0.7054	0.7539	-0.0460	
	(1.7800)	(1.6830)	-(0.7200)	
parking	0.1290	0.1420	0.0476	
	(1.8760)	(2.3170)	(1.0510)	
transportation	-0.1136		0.0193	
	-(0.9530)		(0.2400)	
phone	0.1898	0.2031	0.0010	
	(2.8620)	(3.4570)	(0.0270)	
exc bath	0.1646		0.6730	0.7576
	(0.5060)		(2.2120)	(2.6050)
bathrooms	0.0169		0.0385	0.0408
	(0.3830)		(1.5360)	(1.6530)
kitchen	0.1069		0.1559	
	(0.3130)		(1.1360)	
hot water	0.1295	0.2369	0.0630	
	(1.4590)	(3.9170)	(0.8650)	
AC	0.2096		0.0333	
	(1.0190)		(0.5820)	
material	0.0487		-0.0169	
	(1.2560)		-(0.2730)	
stove	0.1037		0.2381	0.2502
	(0.9950)		(2.1270)	(3.7500)
type bath	0.1228		0.1664	0.0985
	(1.1680)		(1.5510)	(1.7320)
garden	0.0703		0.0145	
	(1.2120)		(0.4100)	
water source	0.0569			
	(0.4030)			
heating	-0.0474		0.0674	0.0750
	-(0.6470)		(2.2900)	(2.7460)
constant	4.4532	4.7829	4.0355	4.1035
	(11.0330)	(40.0890)	(10.2650)	(13.4020)
floor 2			0.1934	0.1973
			(3.3290)	(4.1220)
floor 3			0.1149	0.1348
			(1.4140)	(2.2410)
floor 4			0.2273	0.2661
			(1.6670)	(5.9270)
floor 5			0.2242	0.2670
			(1.5610)	(4.7780)
elevator			-0.0082	
			-(0.0650)	
doorkeeper			0.0372	
			(0.7200)	
F	(22.178) = 7.91	(7.193) = 17.66	(26.316) = 8.88	(11.331) = 14.94
R2	0.5008	0.4739	0.4142	0.4002
n	201	201	343	343

Source: Authors' calculations based on the ENGH.
Note: t-statistics in parenthesis

Table 3 a): Quantile regression estimation. Model for houses

	Coefficient	Std. Error	Conf. Int. Low	Conf. Int. Up.
0.1				
bedrooms	0.0230	0.0740	-0.1228	0.1688
phone	0.1771	0.1236	-0.0663	0.4206
water	0.2001	0.1244	-0.0450	0.4453
security	0.5826	0.6108	-0.6210	1.7862
parking	0.1771	0.1339	-0.0868	0.4411
hot water	0.4308	0.1863	0.0636	0.7980
e.doorbell	0.2053	0.1385	-0.0676	0.4783
intercept	4.6214	0.2275	4.1731	5.0697
0.25				
bedrooms	0.1385	0.0452	0.0495	0.2276
phone	0.1551	0.0733	0.0106	0.2996
water	0.2338	0.0902	0.0561	0.4115
security	0.4868	0.6911	-0.8751	1.8486
parking	0.1657	0.0695	0.0287	0.3028
hot water	0.1551	0.0863	-0.0149	0.3251
e.doorbell	0.1542	0.1581	-0.1575	0.4658
intercept	4.7171	0.1375	4.4461	4.9880
0.5				
bedrooms	0.1069	0.0480	0.0123	0.2015
phone	0.2012	0.0771	0.0493	0.3531
water	0.2624	0.0835	0.0977	0.4270
security	2.0103	0.8860	0.2643	3.7562
parking	0.0677	0.0810	-0.0920	0.2273
hot water	0.2296	0.0769	0.0780	0.3812
e.doorbell	0.1178	0.1438	-0.1656	0.4012
intercept	4.9777	0.1401	4.7017	5.2537
0.75				
bedrooms	0.0967	0.0359	0.0259	0.1675
phone	0.1800	0.0589	0.0640	0.2961
water	0.1933	0.0672	0.0609	0.3257
security	1.7421	0.7571	0.2502	3.2341
parking	0.0944	0.0653	-0.0342	0.2230
hot water	0.2254	0.0570	0.1131	0.3377
e.doorbell	0.2406	0.1749	-0.1040	0.5852
intercept	5.2315	0.1117	5.0114	5.4516
0.9				
bedrooms	0.1542	0.0668	0.0225	0.2858
phone	0.3026	0.1438	0.0192	0.5860
water	0.2364	0.1034	0.0326	0.4402
security	1.5563	0.6248	0.3250	2.7876
parking	0.0531	0.1325	-0.2079	0.3142
hot water	0.1560	0.1180	-0.0766	0.3886
e.doorbell	0.2025	0.2028	-0.1971	0.6022
intercept	5.2395	0.1685	4.9075	5.5715

Source: Authors' calculations based on the ENGH.

Table 3 b): Quantile regression estimation. Model for apartments

	Coefficient	Std. Error	Conf. Int. Low	Conf. Int. Up.
0.1				
bedrooms	0.0558	0.0719	-0.0856	0.1972
bathrooms	0.1691	0.1987	-0.2217	0.5599
heating	0.0000	0.0821	-0.1615	0.1615
exc bath	1.0411	0.4697	0.1173	1.9648
floor 2	0.3152	0.1192	0.0808	0.5496
floor 3	0.0771	0.1473	-0.2127	0.3668
floor 4	0.3152	0.1149	0.0892	0.5412
floor 5	0.1886	0.1425	-0.0916	0.4689
complex	-0.1116	0.1398	-0.3866	0.1635
stove	0.4338	0.2176	0.0057	0.8618
type bath	0.2036	0.1335	-0.0591	0.4663
intercept	3.4295	0.4871	2.4714	4.3876
0.25				
bedrooms	0.1438	0.0193	0.1059	0.1818
bathrooms	0.2513	0.1026	0.0495	0.4531
heating	0.0103	0.0340	-0.0566	0.0773
exc bath	0.2438	0.5122	-0.7636	1.2511
floor 2	0.1438	0.0613	0.0234	0.2643
floor 3	0.0748	0.0759	-0.0744	0.2241
floor 4	0.2774	0.0629	0.1537	0.4010
floor 5	0.1823	0.0836	0.0180	0.3467
complex	-0.2290	0.1097	-0.4447	-0.0133
stove	0.3952	0.1038	0.1910	0.5993
type bath	0.1178	0.0876	-0.0545	0.2900
intercept	4.2642	0.5029	3.2750	5.2535
0.5				
bedrooms	0.1587	0.0224	0.1147	0.2028
bathrooms	0.1878	0.0610	0.0678	0.3078
heating	0.0880	0.0293	0.0304	0.1456
exc bath	0.4028	0.3130	-0.2129	1.0185
floor 2	0.1162	0.0499	0.0180	0.2143
floor 3	0.0788	0.0704	-0.0596	0.2172
floor 4	0.2421	0.0432	0.1571	0.3272
floor 5	0.2421	0.0535	0.1369	0.3473
complex	-0.1452	0.1278	-0.3966	0.1062
stove	0.2513	0.0994	0.0559	0.4467
type bath	0.0092	0.0809	-0.1500	0.1684
intercept	4.4518	0.3254	3.8118	5.0917
0.75				
bedrooms	0.1740	0.0276	0.1197	0.2282
bathrooms	0.1978	0.0796	0.0413	0.3543
heating	0.1137	0.0362	0.0425	0.1850
exc bath	0.2638	0.2486	-0.2251	0.7527
floor 2	0.1445	0.0549	0.0366	0.2524
floor 3	0.1740	0.0566	0.0626	0.2853
floor 4	0.2877	0.0464	0.1963	0.3790
floor 5	0.2719	0.0653	0.1435	0.4004
complex	-0.0488	0.1058	-0.2568	0.1592
stove	0.1740	0.0686	0.0390	0.3089
type bath	0.0281	0.0639	-0.0975	0.1537
intercept	4.6922	0.2473	4.2058	5.1786
0.9				
bedrooms	0.1823	0.0362	0.1112	0.2535
bathrooms	0.2485	0.0870	0.0773	0.4196
heating	0.1178	0.0596	0.0005	0.2351
exc bath	0.2624	0.2404	-0.2106	0.7353
floor 2	0.1054	0.1040	-0.0992	0.3099
floor 3	0.0974	0.1370	-0.1720	0.3668
floor 4	0.2389	0.0915	0.0590	0.4188
floor 5	0.2781	0.1123	0.0573	0.4990
complex	0.0725	0.1697	-0.2613	0.4063
stove	0.2107	0.0793	0.0548	0.3666
type bath	0.0896	0.0895	-0.0865	0.2657
intercept	4.6821	0.2378	4.2143	5.1499

Source: Authors' calculations based on the ENGH.

Table 4: Household income, implicit rent and rent paid

Decile	% of owners	Owners Household income	Imputed rent OLS	Imputed rent QR	% of household income OLS	% of household income QR	Tenants Household income	Rent paid
		(i)						
	(i)	(ii)	(iii)	(iv)	(v)	(vi)	(vii)	(viii)
1	97%	388.8	241.0	205.5	62%	53%	380.8	202.2
2	92%	543.7	292.5	265.4	54%	49%	764.1	257.8
3	90%	693.7	321.9	301.8	46%	44%	813.2	268.0
4	87%	771.5	345.5	341.5	45%	44%	976.8	295.0
5	88%	937.7	374.2	384.6	40%	41%	1035.0	342.3
6	85%	1142.6	399.6	423.8	35%	37%	1194.7	328.9
7	86%	1222.7	407.0	445.6	33%	36%	1238.5	326.8
8	78%	1554.4	454.8	500.7	29%	32%	1339.2	381.3
9	80%	1985.1	479.4	519.7	24%	26%	1711.3	420.6
10	72%	3395.9	542.2	587.3	16%	17%	2290.5	477.6

Source: Authors' calculations based on the ENGH.

Note: Household sorted by household per capita income.

Table 5: Housing characteristics

	Owners		Tenants			
	Mean	Std. Dev.	Mean	Std. Dev.	Difference	|z|-stat
Houses						
bedrooms	3.126	1.110	2.806	0.985	0.320	4.265
phone	0.621	0.485	0.414	0.494	0.208	6.250
water	0.664	0.472	0.746	0.436	-0.081	2.543
security	0.013	0.112	0.017	0.130	-0.005	0.594
parking	0.358	0.479	0.255	0.437	0.102	3.141
hot water	0.624	0.484	0.612	0.488	0.012	0.359
e.doorbell	0.103	0.304	0.078	0.268	0.025	1.240
Apartments						
bedrooms	2.985	1.064	2.296	0.905	0.689	11.138
bathrooms	1.273	0.616	1.057	0.337	0.216	6.409
heating	0.809	0.393	0.641	0.480	0.168	6.556
exc bath	0.992	0.091	0.986	0.116	0.005	0.891
floor 2	0.166	0.373	0.147	0.354	0.020	0.894
floor 3	0.135	0.342	0.120	0.325	0.016	0.764
floor 4	0.282	0.450	0.372	0.484	-0.090	3.246
floor 5	0.215	0.411	0.177	0.382	0.038	1.578
complex	0.112	0.315	0.049	0.217	0.063	3.505
stove	0.948	0.223	0.924	0.266	0.024	1.688
type bath	0.924	0.266	0.929	0.258	-0.005	0.308

Source: Authors' calculations based on the ENGH.

Table 6: Intercept estimations for conditional quantile income distribution of rent

Decile	Logs		Levels		Difference with original intercept			
					pesos		percentage	
	Houses	Apart.	Houses	Apart.	Houses	Apart.	Houses	Apart.
1	4.62	3.43	101.64	30.86	-34.78	-46.40	-25.5%	-60.1%
2	4.72	3.31	112.59	27.34	-23.82	-49.92	-17.5%	-64.6%
3	4.72	4.35	112.50	77.34	-23.91	0.08	-17.5%	0.1%
4	4.86	4.44	128.57	84.38	-7.84	7.12	-5.7%	9.2%
5	4.98	4.50	145.80	90.24	9.39	12.98	6.9%	16.8%
6	5.11	4.52	165.29	91.43	28.88	14.17	21.2%	18.3%
7	5.17	4.73	176.47	113.02	40.06	35.76	29.4%	46.3%
8	5.21	4.80	183.28	121.97	46.87	44.71	34.4%	57.9%
9	5.24	4.68	188.57	108.00	52.16	30.74	38.2%	39.8%
10	5.48	4.42	240.00	83.00	103.59	5.74	75.9%	7.4%
OLS	4.92	4.35	136.41	77.26				

Source: Authors' calculations based on the ENGH.

**Table 7: Distributional effects of implicit rents**

| | Original | Without adjustment for under-report | | | | Adjusted for under-report | | | |
		OLS estimations		QR estimations		OLS estimations		QR estimations	
		Impute rent to credit takers?				Impute rent to credit takers?			
		yes	no	yes	no	yes	no	yes	no
	(i)	(ii)	(iii)	(iv)	(v)	(vi)	(vii)	(viii)	(ix)
Gini	0.445	0.410	0.413	0.418	0.421	0.427	0.428	0.431	0.432
Theil	0.351	0.291	0.295	0.301	0.305	0.320	0.321	0.325	0.326
CV	1.006	0.887	0.893	0.898	0.905	0.947	0.950	0.952	0.954
Atk (e=1)	0.299	0.259	0.258	0.262	0.270	0.279	0.275	0.280	0.285
Atk (e=2)	0.546	0.448	0.454	0.469	0.474	0.486	0.489	0.498	0.500
Atk (e=3)	0.794	0.597	0.605	0.624	0.629	0.650	0.652	0.667	0.668

Source: Authors' calculations based on the ENGH.
Note: CV=coefficient of variation, Atk(e)=Atkinson index with inequality-aversion parameter e.

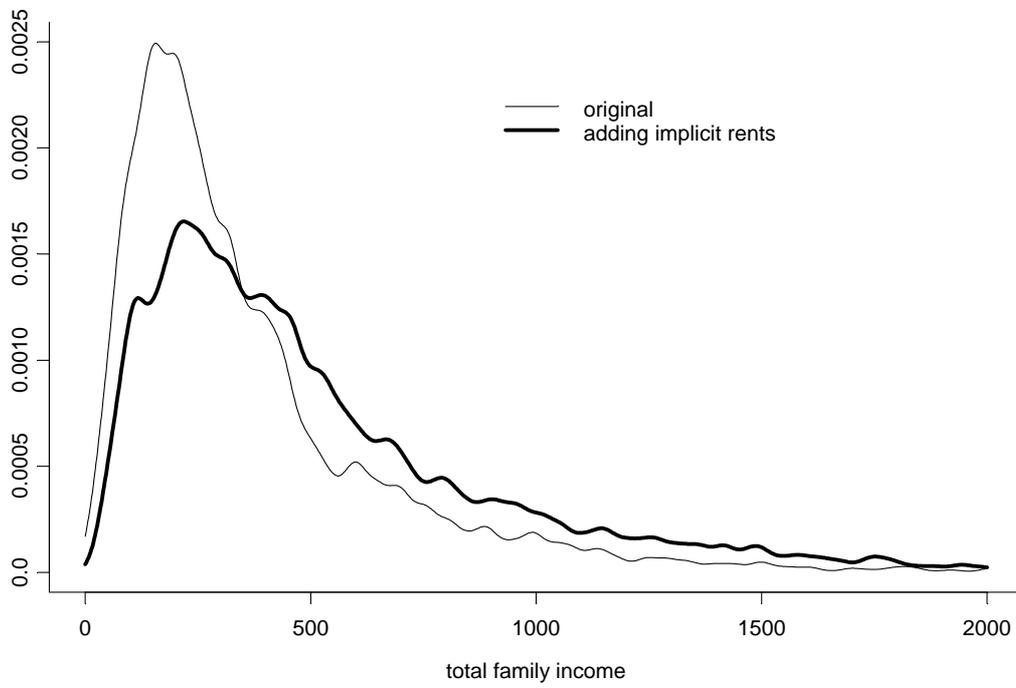**Graph 1: Quantile Regression estimations**

Houses
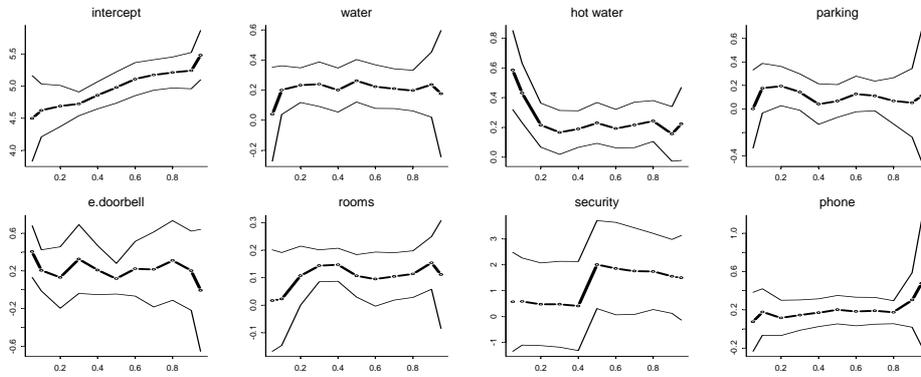
Apartments
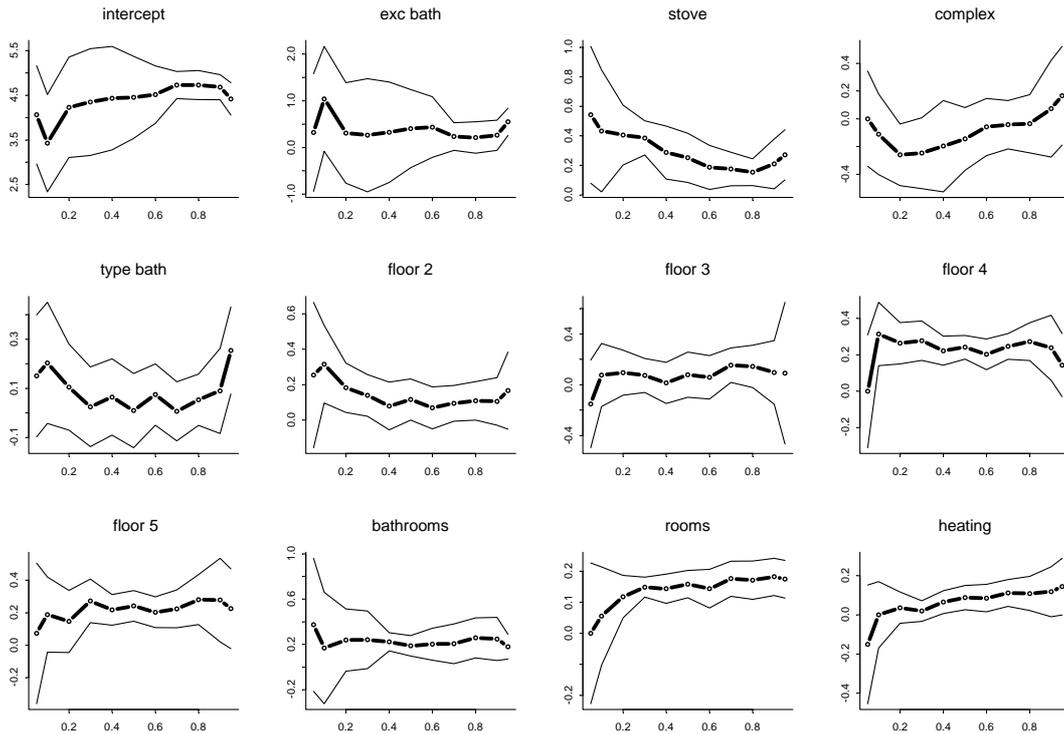
Source: Authors' calculations based on the ENGH.

Figure 2: Comparison of densities of household income

Source: Authors' calculations based on the ENGH.