Journal of Housing Economics 18 (2009) 311–324

Contents lists available at ScienceDirect

Journal of Housing Economics

journal homepage: www.elsevier.com/locate/jhec

The reliability of self-reported home values in a developing country context

Marco Gonzalez-Navarro [a], Climent Quintana-Domeque [b,*]

[a] Department of Economics, Fisher Hall, Princeton University, Princeton, NJ 08544, USA
[b] Departament de Fonaments de l'Anàlisi Econòmica, Universitat d'Alacant, Sant Vicent del Raspeig, Alacant 03690, Spain

ARTICLE INFO

Article history:
Received 14 October 2008
Available online 12 August 2009

JEL classification:
R21
R31
R14

Keywords:
Bias
Inaccuracy
Housing prices
Owners' estimates
Appraised values

ABSTRACT

This paper provides an analysis, in the context of a developing country, of the reliability of homeowners' estimates of the value of their houses, as obtained through a household survey. We show that non-response to the home value question by the owner is uncorrelated with the appraised value of the house and other demographic characteristics of the respondent. We also document that homeowners with long tenure largely overestimate the value of their home. Moreover, both the bias and the lack of precision in homeowners' estimates are correlated with tenure, but not with socioeconomic characteristics. However, we also show that self-reported home values from short-tenure homeowners can be used to obtain unbiased and precise estimates of the average house value at the census tract level.

1. Introduction

1.1. Why should we care about the reliability of self-reported home values in developing countries?

Economists and policymakers often are in need of a trustworthy measure of the wealth of families. For homeowners, any such measure would not be complete without an assessment of the market value of the family's house. Housing wealth has been shown to be a key variable in decisions such as retirement (see, for example, Lusardi and Mitchell, 2007), consumption (see, for example, Campbell and Cocco, 2007), savings (see, for example, Juster et al., 2005; Klyuev and Mills, 2006), and the debt composition of the household (see, for example, Disney et al., 2006).

A family's residence is often its single most valuable asset, particularly in developing countries where financial assets are not very common. By obtaining such a measure, we would be gathering important information about the overall financial position of consumers and their wellbeing, which would be useful in a myriad of applications. For example, information on home values is crucial to policymakers interested in implementing well-informed public policies in such areas as taxation and infrastructure provision. Moreover, if household economic behavior is based on perceived rather than actual wealth, this data could help to determine whether people systematically underestimate or overestimate their actual wealth endowments, and thus whether their economic decisions are suboptimal in an intertemporal context. For example, Agarwal (2007) finds that people who underestimate the value of their home are more likely to prepay their loans, while people who overestimate it are more likely to default on their loans.

In many developing countries, homeowner estimates of property market price obtained through national household surveys may be the most convenient and reliable way of tracking home value. When declared transactions

* Corresponding author. Fax: +34 965 903 898.
E-mail address: climent@ua.es (C. Quintana-Domeque).

1051-1377/$ - see front matter © 2009 Elsevier Inc. All rights reserved.
doi:10.1016/j.jhe.2009.07.013

prices for homes are grossly underreported to the authority in order to evade taxes, the repeat sales methods that are common in the US (like the S&P Case-Schiller index) to study home market dynamics become infeasible. In this paper, we assess the reliability of survey obtained homeowner estimates of house value in a developing country context (Acayucan, Mexico) and argue that if attention is limited to estimates by short-tenure owners, the average home value estimates are reasonably unbiased and precise.

1.2. Literature review on the reliability of self-reported home values

In the United States, all the major household surveys—the decennial Census, the Panel Study of Income Dynamics, the American Housing Survey, and the Survey of Consumer Finances—ask a question like "*What is the value of this property; that is, how much do you think this property would sell for if it were for sale?*" Also, a quick look at the World Bank's Living Standards Measurement Study (LSMS) Survey Finder indicates that 79 LSMS household surveys contain questions on household ownership and value. The main argument favoring the use of the question on home valuation is its ease of collection and wide availability. It is thus crucial to assess the reliability of self-reported home valuations.

Several papers have sought to investigate the reliability of the answers to home valuation questions. In Kish and Lansing (1954), homeowners were asked to estimate the market value of their houses. Consequently, estimates for the same homes were made by professional appraisers. The main finding was that the average bias in people's estimates was around zero. That is, although individuals' estimates could be quite different from the appraised values, the errors seemed to be mean zero and would cancel out on average. This was an important finding justifying the use of the question in large surveys. Regarding the accuracy of the estimates, the researchers found that around 37% of the responses were within an interval of ±10% of the appraiser's estimate. When the researchers focused on different subgroups, their results showed that new homeowners made the most precise estimates of their house value, but there was no increase in accuracy if the respondent was exclusively the household head, if the respondent had more education, or if the agent was able to enter the property during the appraisal.

Using the same methods and similar data as Kish and Lansing, Kain and Quigley (1972) confirmed that errors were largely offsetting, but that they were correlated with the socioeconomic characteristics of the respondent. More education was associated with a smaller positive bias in the homeowner's estimate. This research also inquired into the determinants of non-response to the question, and found that those with higher income, education, and lower tenure in the home were more likely to answer—that is, give an estimate of the value of the house, conditional on participating in the survey.

Because the ideal estimate of the market price of a house is the sale price at the time of purchase, some studies have compared sales data for recently transacted homes with owners' estimates. This method is used by Goodman and Ittner (1992). They compare owners' estimates with subsequent sales prices of the same property using the 1985 and 1987 American Housing Survey. They find that the average US homeowner overestimates the value of his house by 6% over what it actually sells for and that the average absolute error is around 14%. The error is largely unrelated to the characteristics of the owner, the house, or the local market. One concern with Goodman and Ittner's study, though, is that the estimates used were obtained from people who in all likelihood had been planning to sell and were probably gathering information on the value of their property before it was actually sold. This makes soon-to-be sellers markedly different from the rest of the population.

Other studies that use the American Housing Survey to compare transaction prices with owner valuations include: DiPasquale and Somerville (1995) and Kiel and Zabel (1999). Consistent with Goodman and Ittner (1992), both studies find that, on average, owners' valuations exceed transaction prices. Additionally, they find that longer tenured homeowners have a lower estimated value of their property, *ceteris paribus*.

Another approach in some studies is to use tax assessments and compare them to people's self-valuations, as in David (1968). The obvious problem with using tax assessments is that they are generated by local governments who charge property taxes. If the assessments are generated from what citizens report as the sale price, as in the case of Mexico, then there is an incentive compatibility problem, usually resulting in gross underestimations of the value of homes. If the data are obtained from independent assessments by government officials, then there is a tendency to update values infrequently because the higher property tax payments generated are unpopular. This results in a flawed impression of housing value.

Overall, these studies have found that, on average, owners tend to overestimate the value of their homes by around 5%, that this overvaluation is unrelated to owner and home characteristics other than the length of tenure in the home, and that surveys can be reliably used to obtain reasonable estimates of home valuation at a very low cost. However, given that the existing literature has focused on developed countries (mainly the United States), it may not provide a good guide to results in developing countries, where housing market conditions may be different. Indeed, housing market conditions in developing countries differ from those in the United States in several dimensions: most importantly, access to land, construction, and financing, all of which generate a lower level of information about the distribution of house prices.

1.3. Housing market conditions in developing countries

In developed countries, access to land occurs mainly through purchases, while in poor countries it is not uncommon for a substantial proportion of urban growth to occur through squatting, especially for the poor. Inhabitants of cities can organize themselves and invade government lands, protected areas, and even private property. By the time tribunals establish the illegality of such actions, some politicians will find it attractive to provide protection and

services to the squatters in exchange for votes and political support, rather than removing them from the ill-gained land. Local governments commonly engage in expropriation of land that is handed out later to political constituents. Those who acquire their property under such conditions have a harder time finding out what the monetary value of the property is, given that they did not initially pay for it. And, in many cases, the property cannot be sold easily for lack of a valid title.

The second reason that housing markets are different in developing countries is that a much larger proportion of the housing stock tends to be self-built. The presence of self-built houses, instead of developments by specialized construction companies, occurs because the financial system is underdeveloped, and mortgages are either non-existent or very expensive. Families thus acquire a home in these conditions by building it themselves over long periods of time. The lack of developed mortgage markets can thus force families into inefficient construction methods because the house is built slowly over time. For example, the family may buy first a tin roof, and later on replace it with cement. Credit constraints can thus generate unnecessarily high construction costs. If people estimate their home's worth as the sum of the expenses incurred in building it, housing markets where mortgages are non-existent may also present upwardly biased home value estimates as captured by surveys of homeowners.

Another reason why self building can generate a lack of knowledge of the current market value of the property is by losing the following powerful information channel. When a group of homes is built by a construction company, there is a great degree of homogeneity of homes in the vicinity, so when one neighbor moves in or out, this will generate information on the current value of the surrounding properties. When housing is self-constructed, this information channel disappears, because the neighboring homes are not a very good proxy for the value of a family's house.

Jimenez (1982) provides one of the few studies for a developing country. Using data from an impoverished neighborhood in the Philippines, he finds that the mean value owner and appraiser estimates are not statistically distinguishable. His Philippine sample does well when compared with Kain and Quigley's results for St. Louis differences in average valuations. However, for individual estimates, his results are rather different. The average of the absolute value of the differences between owner and appraiser valuations is approximately 55% of the mean appraised value, while the comparable figure for Kain and Quigley is approximately 20%. Jimenez also inquires whether the role of using owner versus appraiser valuation changes the relationship between housing prices and their characteristics. The overall results appear to be similar when a comparison is made between the coefficients of equations with different dependent variables (owner's valuation and appraised valuation). The signs and magnitudes are roughly of the same order of magnitude.

1.4. Contribution of this study

This paper investigates the reliability of homeowners' estimates of the value of their houses in the context of a developing country. Self-estimates of home value in such countries may be simply unreliable with thin housing markets and widespread non-market access to land and housing, including squatting. The research presented here complements and extends upon Jimenez's work in at least three different ways. First, we try to determine how reliable self-valuations are, but also look at the magnitude of the bias, the accuracy of people's answers, and their determinants. Second, we explore whether the relationship between housing prices and housing characteristics depends on the housing price estimate (owner's estimate versus appraised value), with the advantage of having two independent measures for several housing characteristics. Finally, we show how our results can be used to generate estimates of average home values at the neighborhood or census tract levels.

We use data from a household survey and appraisals for the same homes done by a real estate agent. The survey was applied from mid-February to mid-March 2006 to approximately 1200 dwellings in the outskirts of the city of Acayucan, Veracruz in central Mexico. The head of the household or the spouse of the head of the household was interviewed, with the main respondent providing information about the rest of the family. A detailed description of the survey can be found in Gonzalez-Navarro and Quintana-Domeque (2007). The city serves as the local trading center in an agricultural and livestock-producing region. Its population is slightly less than 50,000 and incomes are somewhat lower than the Mexican average.[1]

We must emphasize that the sample used for this study is neither representative of developing countries nor of Mexico as a whole. However, it is a representative random sample of the outskirts of the city of Acayucan. The sampled areas of the city corresponded to the poorer parts of town, where streets are not paved, and many homes are lacking such important services as sewerage, running water, and indoor plumbing.

1.5. Main results of this study

Our findings suggest that the ability of homeowners to estimate own-residence value is mainly a function of how short their tenure is. Owners with long tenure overestimate by a large amount the value of their home (in some cases by more than 200%). However, families with short tenure have reasonably accurate and unbiased estimates of the value of their home, similar to what is found in the US literature. Our regression analysis, where we control for the discrepancy in the lot-size estimate between the owner and the real estate agent, several socioeconomic characteristics, housing characteristics, and neighborhood effects, seems to confirm that both the bias and lack of accuracy in owners' self-valuation reports are driven by tenure (years living in the house). Neither the bias nor the lack of accuracy in owners' self-valuation reports is

[1] According to the 2000 Census, 14% of income earners made less than 1 minimum wage, while 33% of Acayucans were in this category. Fifty percent of Mexicans earned at most two minimum wages, while 70% of Acayucans earned in this income bracket.

correlated with socioeconomic characteristics. Interestingly, we also find that item non-response is uncorrelated with the appraised value of the house. This suggests that unbiased estimates of the average value of a group of homes can be obtained through household surveys without being worried about selectivity concerns. Another interesting result is that inhabitants of neighborhoods with relatively homogenous homes built by a construction company appear to make unbiased and very precise home-value estimates. Finally, using hedonic regressions, we find that homeowners have trouble determining the plot size of their home in terms of square meters. This is evidenced by the fact that the owner's estimate of the size of the lot is uncorrelated with the appraiser's valuation.

Our paper is organized as follows. Section 2 provides a description of the data used in the paper. Section 3 presents the results. Section 4 illustrates how these results could be used to generate estimates of mean home prices at the census tract level. In Section 5, the reliability of the real estate agent's appraisals is investigated. Section 6 provides the conclusion.

2. Data

2.1. Data sources

The owner-estimated home valuations come from the Acayucan Standards of Living Household Survey (ASLS) 2006 (Gonzalez-Navarro and Quintana-Domeque, 2007). This household survey contains detailed information about demographic, income, education, and housing characteristics of the sampled population. Among the housing questions of the ASLS, one asked respondents to estimate the market value of their houses. Specifically, the question was phrased: "*Approximately how much money do you think this house would sell for nowadays?*". The interviewed families mostly lived in small, single floor homes, with a wall-delimited lot. Of the respondents, 51% were spouses of the household head and 45% were the household heads themselves. The survey had a response rate of 94%.

The other source of data in the study was the housing value assessments produced by a real estate agent. Having only one real estate agent perform all the assessments has the positive feature that heterogeneity of assessment practices, which require a lot of subjective decision-making, was minimized. The real estate agent was asked to visit one out of every two successfully interviewed homes and to assess the market value of the house. The appraisals then were matched to the interviews, with a success rate of almost 90%. The assessments were performed within two months of the household interviews.

Our study, as well as previous ones in this literature, did not ask the appraiser to enter a home to perform the valuation. The main reason for not doing so is that the homeowners would find such a request extremely intrusive. Given that this study is part of a larger ongoing research program which involves multiple survey rounds in the same homes, we decided against asking the appraiser to attempt entering the houses with the objective of minimizing future survey non-response. Another discarded strategy was asking the appraiser to accompany the survey team and perform his valuation concurrently with the interview. This was not done because in the pilot and in the experience of the survey company, most interviews would be performed at the entrance of the house, and few homeowners would actually allow the survey worker inside the house.

One important advantage of our procedure is that we are the first to investigate what the appraised home values are for respondents that did not provide an answer to the question in the survey. Hence, unlike any previous study, we can look at whether non-response to the home valuation question by the owner is related to home value as measured by the appraiser. Given that the professional appraisals were preformed within a 2-month period after the survey, there are no concerns that house price inflation or volatility are relevant for our results.

2.2. Socioeconomic and housing market characteristics

Table 1 contains some of the main socioeconomic characteristics of the individuals surveyed. It is divided into three panels. The top panel shows descriptive statistics for the whole sample: 53% of the persons in the sample are female, with an average age of 28 years, and an average education level of 7 years. The mean household size is four persons, and their monthly per capita expenditure is approximately $86 (2006 US dollars) using market exchange rate, or $131 in PPP terms. For comparison purposes, the middle panel has similar variables for Mexican cities of the same size category. The surveyed sample is slightly older (2 years), somewhat more educated, and has an almost identical mean household size to that of similar sized cities in Mexico. Regarding the respondents, 78%

Table 1
Demographic and socioeconomic characteristics of survey sample, similar sized cities, and respondents.

	Observed[a]	Mean	Median	SD
Survey sample				
Female	4972	0.53	—	0.50
Age	4967	27.85	24	19.59
Years of schooling	3955	7.24	6	4.27
Household size	1239	4.01	4	1.80
Per capita expenditure (2006 US dollars)	1211	86.82	68.18	75.06
Similar sized cities[b]				
Female	529,209	0.51	—	0.50
Age	529,209	25.81	22	19.22
Years of schooling	529,209	6.45	6	4.43
Household size	125,768	4.29	4	2.07
Respondents[c]				
Homeowner	1239	0.70	—	0.46
Female	1239	0.78	—	0.42
Age	1236	42.44	40	14.90
Years of schooling	1018	7.85	7	4.26

[a] The numbers of observations for the household size and per capita expenditure variables refer to the number of dwellings, unlike the other variables, which refer to individuals.
[b] Data for similar sized cities from IPUMSI micro data, corresponding to the 2000 Census.
[c] Includes all respondents to the questionnaire, regardless of item non-response.

were female and the average age was 42 years. Respondents had around half a year more education than those in the whole sample.

The survey fieldwork permitted us a deeper knowledge of the housing market in Acayucan than the survey instrument by itself would have provided. Already built homes were found to be acquired mostly in traditional ways: purchase, inheritance, or donation within the family. However, some of the families living in self-constructed homes, especially the poorer ones, had a peculiar way of acquiring the land. Frequently, poor people obtained their lots for a very modest payment (around 10% or 15% of the market price) from the municipality, which owns several tracts of land in the surrounding parts of the city. Many of the households in the sample that declared not having a property title had not formalized the protocol of transfer of property from the municipality. The other way that families acquired land in the outer parts of this city was through direct purchases from landowners who subdivided large plots of land for sale. In either case, families started out with no public services. The transacted plots usually were marked pieces of land with space left for streets that would eventually get electrification, pavement, and water services.

We can provide two measures of the intensity of housing market transactions. First, looking at the Public Registry of Property in Acayucan, we see that over the course of 2005, approximately 2% of homes were registered as having been sold in Acayucan. This can be thought of as a minimum estimate of the proportion of transacted homes, because a fraction of those transactions that took place in 2005 had not been registered yet: new proprietors may decide to occupy the house first and pay the transaction taxes, which are necessary for the property to become publicly registered under the new owner's name, later on. Second, our household survey provides information on the proportion of homes that were transacted during the year. In the sample, the percentage of owned-houses in which the respondent has been living for 1 year or less is 4.75%. This can be thought of as an upper bound on the annual percentage of houses that are sold, because some of these homes are being "entrusted" to the occupants by friends or family. This intensity of transaction, between 2% and 4.75%, compares to that of the United States of 5% per year (Goodman and Ittner, 1992).

In the completed surveys, the item response rate for the owners' estimate of the house value was approximately 74%. Interestingly, we cannot reject the hypothesis that the average appraised value is equal for respondents who did and did not provide a house value estimate. Indeed, Table 2 shows that the probability of response is not related to the appraised home value conditional on age, sex of the respondent, household head status (i.e., whether the respondent is the head of the household or not), having a property title, tenure (years living in the house) and 16 neighborhood dummies. Additionally, age, sex, household head status, having a property title and tenure are not related to the probability of response. This suggests homeowners that do not provide an estimate of house value are a random subset of our sample. This is an important finding if home owner valuations are to be used to estimate average home prices in a locality.

Table 2
Determinants of response to the home valuation question. Probit estimates. Marginal effects evaluated at the mean are reported.

	(1)	(2)	(3)	(4)	(5)
Appraised home value (P_a)	0.000	−0.000	−0.000	−0.000	−0.000
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Age	—	−0.001	−0.000	−0.000	−0.002
		(0.002)	(0.002)	(0.002)	(0.003)
Female	—	−0.059	−0.083	−0.085	−0.093
		(0.064)	(0.073)	(0.072)	(0.069)
Household head	—	—	−0.052	−0.060	−0.063
			(0.066)	(0.065)	(0.065)
Property title	—	—	—	−0.080	−0.083
				(0.071)	(0.070)
Tenure	—	—	—	—	0.003
					(0.003)
Neighborhood dummies?	No	Yes	Yes	Yes	Yes
p-Value F-test	—	0.00	0.00	0.00	0.00
Pseudo-R^2	0.00	0.08	0.08	0.08	0.09
Sample size	363	359	359	357	350

Note. Clustered robust standard errors are reported in parentheses (52 clusters). Neighborhood dummies: 16 dummies based on the first three digits of the AGEB (Mexican geographical unit) classification system.

2.3. Characteristics of owner-occupied homes reported by the owner and the real estate agent

Table 3 presents some relevant characteristics of the houses that are the focus of this study: owner-occupied homes.[2] In terms of the size of the lot, the estimates are 240 and 130 m^2 for the owners and the appraiser, respectively. The difference in assessed lot size by the appraiser and the respondents is large and significant. The average house in the sample has 2.5 rooms,[3] and surprisingly, only 63% of the dwellings have a bathroom inside the house. Cement roofing is not the norm in the sample: 60% of dwellings have a roof made out of metal sheets, asbestos, or palm leaves. Further, 12% of homes did not report having a property title. We also realize that there is a large discrepancy in the average value obtained from the owners and the real estate agent. The average owner's estimated home value is $19,948, while the average appraisal is only $12,123. The median difference, although much smaller ($1545), and the mean log difference, are also significantly different for both measures.

In the paper, we assume that the real estate agent valuation is very close to the market value of the house. We think that it is reasonable to interpret the discrepancy between home values obtained from the owners and the real estate agent as originating from homeowner's misperceptions about home market value. There are several reasons justifying such an interpretation. First, the appraiser has a more sensible estimate of the lot size (one of the most important determinants of home value according to the real estate agent) than the owners. Second, the real estate

[2] Although the appraiser performed around 600 valuations, 26% were not used due to a lack of response from homeowners (item non-response), and another 30% were lost due to the respondents not being homeowners.

[3] Rooms here refer to the number of rooms in the house, excluding the kitchen, for any use, not necessarily for sleeping.

Table 3
Characteristics of owner-occupied homes.

	Observed	Mean	SD	Median	Mean log
Self-assessed measures					
Home value ($ US)	267	19,948	35,595	9091	9.30
Lot-size estimate (m^2)	260	240	176	200	5.27
Property title	266	0.88	0.33	—	—
Tenure	263	17.77	13.65	—	—
Number of rooms[a]	267	2.51	1.14	—	—
Bathroom within house	267	0.63	0.48	—	—
Sewerage	267	0.87	0.34	—	—
Cement walls	266	0.93	0.25	—	—
Cement roofing	267	0.40	0.49	—	—
Cement floor	267	0.95	0.22	—	—
Appraiser measures					
Home value ($ US)	267	12,123	10,273	7546	9.13
Lot-size estimate (m^2)	267	130	52	120	4.79
Bathroom within house	268	0.91	0.28	—	—
Sewerage	268	0.99	0.12	—	—
Cement walls	267	0.96	0.20	—	—
Cement roofing	268	0.40	0.49	—	—
Cement floor	268	0.96	0.20	—	—

	Corr.	Difference of means	Difference of mean logs
Home value ($ US)	0.16 (0.01)	7825 [0.00]	0.169 [0.00]
Lot-size estimate (m^2)	0.13 (0.04)	−110 [0.00]	0.476 [0.00]
Bathroom within house	0.25 (0.00)	−0.28 [0.00]	—
Sewerage	0.22 (0.00)	−0.12 [0.00]	—
Cement walls	0.47 (0.00)	−0.03 [0.07]	—
Cement roofing	0.53 (0.00)	−0.01 [0.80]	—
Cement floor	0.44 (0.00)	−0.01 [0.41]	—

Note. *p*-Values for the statistical significance of the correlation coefficients are reported in parentheses and *p*-values from paired *t*-tests of equality of means are reported in brackets.

[a] Number of rooms excluding kitchen.

agent seems to infer accurately other housing characteristics. Third, he is likely to be cognizant of the market forces involved in dwellings valuation in squatter communities. Let us now explain these claims.

2.4. First claim: the appraiser has a more sensible estimate of the lot size than the owners

During the fieldwork, we realized that homeowners seemed to have difficulties calculating area. On the other hand, the real estate agent may have a faulty impression of the lot size of the property. However, the latter seems unlikely. Although the real estate agent did not have access to the interior of the home, he could observe the entire lot and knew where the property lines fell, given that homes are delimited by low height walls or fences in the survey area. Moreover, his ability, skills, and training in performing home valuation are expected to be much better than those for the owners. Hence, we expect that he provides better estimates of the lot size than the owners. Fortunately, we can provide some evidence supporting such a claim. Note that since lot sizes are correlated within a street block, the between-blocks variance of lot sizes should be larger than the within-blocks variance. When this is tested, the equality of within and between variance cannot be rejected for the self-respondents' lot-size estimate, while it is rejected for the appraiser. That is, for the agent's lot-size estimate, the average between-block deviation is significantly larger than the average within-block deviation. This suggests that the appraiser has a more sensible estimate of the lot size than the residents. Note that all of the appraisals were performed by the same real estate agent. Hence, the previous finding cannot be due to different real estate agents conducting appraisals in different neighborhoods.[4]

2.5. Second claim: the appraiser infers accurately other housing characteristics

Given that the appraiser did not enter the properties, this may suggest that he erred substantially in determining *internal and less visible housing features* (whether the house had a bathroom inside the house or sewage system) but was much better at determining *external and more visible housing features* (whether the home had cement roofing and cement walls), as the means and correlations in the table may suggest. However, a careful look at our data reveals that the real estate agent was very accurate in predicting both types of housing characteristics. Indeed, the real estate agent correctly guessed the presence of a bathroom within the house in 67% of the valuations, 78% for cement roofing, 87% for sewerage, 94% for cement walls, and 95% for cement floor. That is, the real estate agent appears to have accurate information on housing characteristics.

2.6. Third claim: the appraiser is likely to be cognizant of the market forces involved in dwellings valuation in squatter communities

The vitality of the housing market in Acayucan is big enough to provide information on market forces (supply and demand) to the appraiser. Our estimate of the annual intensity of housing transactions in Acayucan is between 2% and 4.75%. This suggests that the professional appraiser is likely to be cognizant of the market forces involved in dwelling valuation in a squatter community, given that his work consists of market value appraisals for bank loans, and property sales advisory services.

Before concluding this section, it is worth arguing that homeowner value estimates may be the only potentially widely available source of house market prices in a developing country. Given that the housing market consists of private transactions between different parties, and that property tax laws generate a tendency to underreport the

[4] Another possible explanation for the large dispersion in lot-size estimates is that survey respondents confuse lot size with construction size. If this were the case, the data would show a positive correlation between appraised building size and the error measures. This was not the case, so we discard this hypothesis in our data.

transaction price to the authorities, home value estimates like those used in the US that depend on reported transaction prices are simply not reliable sources of market value in many developing countries. In Section 5, we study a subset of recently sold single-family properties in Acayucan and obtain market value measures from the tax authorities, the professional appraisal, the homeowner declared transaction price, and current valuation. These measures were obtained with the objective of assessing the reliability of professional appraisals in capturing market transaction prices.

3. Empirical analysis

3.1. The relationship between owners' estimates and appraised estimates

The analysis presented below only uses data from households in which the owner lives with the interviewed family, eliminating answers from families that rent or borrow their homes. Table 4 reports the correlation coefficients between the owner's estimate and the appraised value for the owners' sample and for some interesting subsamples. For the owners' sample, the correlation between the logs of these two measures is 0.42. The correlation increases slightly to 0.43 when the sample is restricted to owner-occupied homes with a property title. One interesting phenomenon is that lower tenure in the house is associated with a larger correlation between the two measures. The correlation increases to 0.61 for owners with less than 2 years of tenure. For those with longer tenure (3 years or more) the correlation is 0.40. This finding is consistent with the standard explanation given in the housing literature: recent movers have more accurate information about the value of their home, since their reports are likely to be close approximations to the market value at the time of purchase.

In Table 5, we tabulated a partition of the owner's estimates as a proportion of the appraiser's estimates. Two main features of this table stand out. First, only 11% of the estimates by respondents are within 10% of the appraiser's estimates. Second, the two measures do not seem to be closely related: 25% of respondents think their home is worth at most 70% of its appraised value. Meanwhile, 35% of respondents think their home is worth at least 150% of its appraised value.

These findings contrast with the existing evidence for developed countries. For example, Kish and Lansing (1954) and Kain and Quigley (1972), respectively, report that only 6% and 12% of the respondents in their sample assessed their home value to be less than 70% of the appraised one. Meanwhile, 9% and 8% estimated it to be larger than 150% of the appraised value, and 26% and 37% had an estimate within 90–109% of the real estate agent's estimate. This suggests that relative to the United States, our survey respondents' are less precise in their estimate of home value. However, our results are in line with those for developing countries. In Jimenez (1982), 25% of respondents think their home is worth at most 70% of the real estate agent's estimate, 41% estimated it to be larger than 150% of the appraised value, and 8% had an estimate within 90–109% of the real estate agent's appraised value.

We now turn to a description of the bias and inaccuracy of owners' estimates. Table 6 shows the average degree of error and lack of precision in owners' estimates for the different subsamples. As in the previous literature, our results are shown for different measures of bias (the difference between owner's and appraisal's home value estimates, and the percentage difference in terms of appraisal's estimate) and inaccuracy (the absolute difference and the absolute percentage difference).

Using all owners in the sample (first column) the average difference between owner's estimate and appraised value is around $7800; this means that owners tend to largely overestimate the value of their home. The mean percentage difference is close to 124% of the appraised value. In terms of inaccuracy or lack of precision, the mean absolute difference is approximately $13,500. This is evidence of how different the appraiser's estimates are from those of the homeowners': on average, the people in the sample have an unrealistically high estimate of the value of their home. These results contrast with the available evidence for the US and the Philippines. Both Kain and Quigley (1972) in St. Louis and Jimenez (1982) in the Philippines report a mean percentage difference of less than 0.5%. In terms of precision, we also find very different

Table 5
Distribution of owners' estimates as a percentage of appraised value.

%	$(P_s/P_a) * 100$ (%)
<70	25
70–89	11
90–109	11
110–129	12
130–149	6
⩾150	35

Note. Authors' calculations.

Table 4
Correlation of owners' estimates and appraisals.

	Owners' sample	Owners' sample and property title	Owners' sample and tenure ⩽ 1 year	Owners' sample and tenure ⩽ 2 years	Owners' sample and tenure > 2 years
$Corr(P_a, P_s)$	0.16	0.20	0.64	0.66	0.15
p-Value	0.01	0.00	0.03	0.00	0.02
$Corr(\ln(P_a), \ln(P_s))$	0.42	0.43	0.49	0.61	0.40
p-Value	0.00	0.00	0.11	0.00	0.00
Sample size	267	233	12	23	244

Note. P_a and P_s refer to the professionally appraised home value and the self-assessed home value, respectively.

Table 6
Owners' average bias and inaccuracy.

	Owners' sample	Appraised value > $10,000	Tenure ⩽ 1 year	Tenure ⩽ 2 years	Non-self-built neighborhood
Self-assessed home value (P_s)	19,948	25,918	14,848	14,625	26,903
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
Appraised home value (P_a)	12,123	22,184	14,189	12,972	25,039
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
Error	7825	3734	659	1652	1864
$(P_s - P_a)$	(0.00)	(0.36)	(0.83)	(0.39)	(0.50)
Percentage error	1.24	0.25	0.21	0.36	0.11
$(P_s - P_a)/P_a$	(0.00)	(0.17)	(0.42)	(0.11)	(0.32)
Absolute error	13,517	16,279	7371	6462	8019
$\lvert P_s - P_a \rvert$	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
Absolute percentage error	1.59	0.75	0.58	0.69	0.33
$\lvert (P_s - P_a)/P_a \rvert$	(0.00)	(0.00)	(0.01)	(0.00)	(0.00)
Clusters	52	38	11	16	1
Sample size	267	101	12	23	14

Note. Clustered robust *p*-values are reported in parentheses (52 clusters), except in column (5), which corresponds to cluster number 11.

results. In our sample, the absolute percentage difference is estimated to be more than 150%, while this is approximately 55% in Jimenez, and 20% in Kain and Quigley. Note that our sample size (267 observations) is higher than the ones used in Jimenez (96 observations) and Kain and Quigley (113 observations).

Fig. 1 pinpoints one source of error in homeowners' valuations. In the figure, people's valuations are divided into deciles according to the real estate agent's appraised value of the house. For each of those deciles, the mean error is calculated and plotted in the figure. The graph suggests that people who live in homes with appraised market value in excess of $10,000 have a much better idea of how much their home is worth. The much noisier observations are clustered mostly where valuations are less than $10,000. This suggests that people who live in very low-value homes tend to grossly overestimate the value of their home, while those who live in more expensive homes have much smaller biases. Although this is quite different from what was documented in Kish and Lansing (1954), who find a close-to-zero bias for every group of appraised value, Martinelli and Parker (2009) point out that a small amount of misreporting may be due to embarrassment about the low value of the house.

Documentation of the bias and inaccuracy for owners with a high appraised value (in excess of $10,000), is presented in the second column of Table 6. The error is much smaller, as the graph suggested. However, learning that people who live in low-value homes have a very large upward bias in own-home-value estimates is scarcely useful for the purpose of determining whether household surveys can be used to infer home values. If both groups declare large home values, then it is impossible to distinguish each group without professional appraisals. However, the

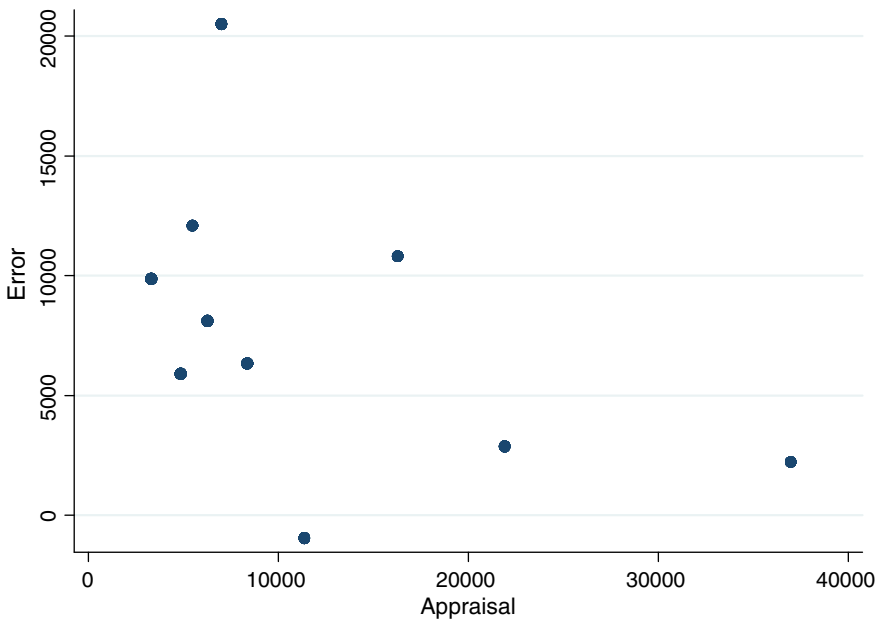

Fig. 1. Mean bias by decile of appraised value.

analysis that follows will suggest that years of tenure is the only significant predictor of bias and lack of precision in home value estimates.

The error and inaccuracy for short-tenure owners (less than either 1 or 2 years of tenure) are reported in the third and fourth columns of Table 6. It turns out that if the analysis is restricted to owners with a short tenure, the mean bias of the estimates is not statistically different than zero; the same holds for mean percentage error. The absolute percentage error also is reduced by over 50%.

In Section 1, we mentioned that one of the differences between developed and non-developed housing markets was the lack of home construction by specialized companies, and that this generated less information about the distribution of home prices for people inhabiting those homes. The last column of the table, column five, isolates the set of homes that were not self-built and estimates the bias and precision of homeowners' estimates. For this subgroup, the mean error and the mean percentage error are not statistically different from zero, while the absolute percentage error is the smallest of the five groups presented in the table (33%). Although we are not aware of any study measuring the magnitude of the inefficiency in construction costs generated by credit constraints, owners of self-built homes may provide upwardly biased estimates because building over time is more expensive than doing it in one step by a specialized company.

The findings in the second, third, and fourth columns of the table are reconciled by the observation that owners with many years in the same home have lower appraised home values. However, longer tenure is also correlated with higher self-assessed home value.

The evidence presented up to now suggests two hypotheses: recent homeowners provide an unbiased estimate of home value, and/or high-value homeowners provide unbiased estimates of home value. In Table 7, we test these two hypotheses. The table is divided into four panels, one for each of the bias and accuracy measures. The top-left panel, corresponding to the error measure $(P_s - P_a)$, presents the predicted error for each of the four analyzed cases: long-tenure or short-tenure and high-value or low-value house. The p-value from a joint test of significance for the coefficients is reported in parentheses. As was discussed above, the panel indicates that short-tenure homeowners have zero mean bias. High-value homeowners also have mean zero bias. The fact that there is mean zero bias for short-tenure and low-value homeowners is important. It confirms that an unbiased estimate of home value can be obtained using the answers of short-tenure homeowners, and that the misreporting due to embarrassment among families in low-value houses (Martinelli and Parker, 2009) may be less important than misreporting due to lack of information about the housing market, at least in a developing country context like the one in Acayucan.

The top-right panel, focusing on mean percentage error, also finds a mean zero percentage error if the focus group is short-tenure homeowners. Mean absolute error (bottom-left panel) is halved when we focus on short-tenure homeowners. Finally, the bottom-right panel indicates that short-tenure homeowners present a third of the mean absolute percentage error of long-tenure homeowners. Hence, the bottom line seems to be that an unbiased and reasonably precise estimate can be obtained from homeowners with low tenure, even if the value of those homes is small.

3.2. Determinants of individual bias and inaccuracy

The results presented until now suggest that long tenure is responsible for the bias and inaccuracy in home owners estimates. This subsection explores the role of tenure in a regression setting. Our purpose is to obtain the effect of tenure on the bias and the lack of precision of homeowners' estimates conditional on other factors that may be related to tenure.

Conversations with the real estate agent suggested that one of the most important determinants of the value of the home is the lot size. We already noted the discrepancy between the estimates of the lot size reported by the owners and appraised by the real estate agent. This discrepancy may be related to both tenure and the discrepancy between the valuations, so it is important to include the discrepancy in lot size into the regressions. Although perhaps this is the most important channel of discrepancy in home valuations that can be related to tenure, some other factors may play a role. To account for the influence of these other factors, we add several other controls: socioeconomic

Table 7
Tests of bias and inaccuracy according to tenure and appraised home value. OLS estimates.

Model: Dependent variable = $\alpha + \beta_1 I_{\text{long tenure}} + \beta_2 I_{\text{high value}} + \beta_3 I_{\text{long tenure}} I_{\text{high value}}$

	Dependent variable: error $(P_s - P_a)$			Dependent variable: percentage error $(P_s - P_a)/P_a$		
	Short tenure	Long tenure		Short tenure	Long tenure	
Low value	3649 (0.19)	10,566 (0.00)	10,273 (0.00)	0.46 (0.19)	1.91 (0.00)	1.84 (0.00)
High value	−3528 (0.53)	4247 (0.32)	3867 (0.34)	−0.14 (0.63)	0.28 (0.14)	0.26 (0.15)
	659 (0.82)	8163 (0.01)		0.21 (0.39)	2.29 (0.00)	
	Dependent variable: absolute error $\lvert P_s - P_a \rvert$			Dependent variable: absolute percentage error $\lvert (P_s - P_a)/P_a \rvert$		
	Short tenure	Long tenure		Short tenure	Long tenure	
Low value	4506 (0.09)	12,126 (0.00)	11,803 (0.00)	0.59 (0.07)	2.16 (0.00)	2.10 (0.00)
High value	11,381 (0.00)	16,541 (0.00)	16,289 (0.00)	0.55 (0.00)	0.77 (0.00)	0.76 (0.00)
	7371 (0.00)	13,806 (0.00)		0.58 (0.00)	1.64 (0.00)	

Note. Clustered robust p-values (52 clusters) for test of joint significance of $\alpha = 0$, $\alpha + \beta_1 = 0$, $\alpha + \beta_2 = 0$, and $\alpha + \beta_1 + \beta_2 + \beta_3 = 0$, are reported in parentheses. $I_{\text{long tenure}} = 1$ if tenure is 2 years or more, 0 otherwise, and $I_{\text{high value}} = 1$ if the property is assessed at more than $10,000, 0 otherwise.

Table 8
Determinants of individual bias and inaccuracy. OLS estimates.

Dependent variable	(1) $(P_s - P_a)$	(2) $(P_s - P_a)/P_a$	(3) $\lvert P_s - P_a \rvert$	(4) $\lvert (P_s - P_a)/P_a \rvert$
Discrepancy in lot size[a]	13.19	0.535	9.49	0.537
	(10.84)	(0.334)	(11.94)	(0.373)
Tenure	613**	0.094***	559***	0.089***
	(231)	(0.032)	(207)	(0.030)
Years of schooling	1498	0.158	1410	0.140
	(971)	(0.097)	(941)	(0.093)
Property title	1873	−0.377	−3089	−0.542
	(6462)	(0.923)	(5742)	(0.872)
ln(PCE)	864	−0.128	4100	−0.041
	(4443)	(0.485)	(3810)	(0.464)
Household head	−4602	0.193	−4470	0.193
	(5050)	(0.407)	(4819)	(0.404)
Housing characteristics?	Yes	Yes	Yes	Yes
p-Value *F*-test	0.69	0.21	0.32	0.21
Neighborhood dummies?	Yes	Yes	Yes	Yes
p-Value *F*-test	0.00	0.00	0.00	0.00
R^2	0.32	0.31	0.36	0.31
Sample size	211	211	211	211

Note. All regressions include a constant term. Housing characteristics: number of rooms, cement floor dummy, cements wall dummy, cement roof dummy, and bathroom within the dwelling. Neighborhood dummies: 16 dummies based on the first three digits of the AGEB (Mexican geographical unit) classification system. Clustered robust standard errors are reported in parentheses (52 clusters).

[a] The discrepancy in lot size is defined differently in every column. In the first column as $(m_s^2 - m_a^2)$, in the second column as $(m_s^2 - m_a^2/m_a^2)$, in the third column as $\lvert m_s^2 - m_a^2 \rvert$, and in the fourth column as $\lvert (m_s^2 - m_a^2)/m_a^2 \rvert$.

** Significant at the 5%.

*** Significant at the 1%.

characteristics (namely, years of schooling of the respondent, a dummy for having a property title, and the log of monthly per capita expenditure) and several housing characteristics reported by the owner (an indicator for whether walls are made out of cement, an indicator for whether the roof is made out of cement, an indicator for whether the floor is made out of cement, an indicator for whether the bathroom is inside the dwelling, and the number of rooms). Moreover, we also include neighborhood fixed effects (16 neighborhood dummies based on the first three digits of the AGEB (Mexican census tract unit) classification system) to account for environmental and spatial characteristics related to tenure, bias and lack of precision. Finally, we must also bear in mind that although our sample is restricted to owned dwellings, the respondent is either the head of the household or his spouse, but she is not necessarily the owner of the dwelling. This is important, because the spouse may be less accurate than the household head. We acknowledge that by including an indicator variable that takes on value one if the respondent is the head of the household, and zero if she is the spouse. All the regressions use robust clustered standard errors at the street level to account for the correlation between observations within the same street.[5]

The results in Table 8 show that the unique statistically significant correlate of large bias and less accuracy is long tenure. Surprisingly, neither the discrepancy in the lot-size estimate nor the socioeconomic characteristics of the respondent seem to have an effect on either the error or the lack of precision of homeowners' estimates.

While the story about the role of tenure as the principal reason for the discrepancy between appraisers' and owner's valuation is compelling, it can be just a statistical artifact due to multicollinearity. In other words, if the independent variables (tenure, discrepancy in lot size, years of schooling of the respondent, etc.) are highly correlated, it will be difficult to determine which independent variable is actually predicting the discrepancy. Hence, we need to be aware about the correlation among the independent variables. For example, if we are interested in the assessing the relationship between X_k and the discrepancy level, we need to be aware about the correlation between X_k and the rest of independent variables X_{-k} (all except X_k). If the correlation is high, the variance of the coefficient estimate is being inflated by multicollinearity. The variance inflation factor (VIF) of the coefficient estimate for the variable k can be written as $\text{VIF}_k = \frac{1}{1-R^2_{X_k X_{1-k}}}$, where $R^2_{X_k X_{1-k}}$ is the R^2 of a linear regression of X_k on the rest of independent variables X_{-k}. The square root of the VIF tells us how much larger the standard error is, compared with what it would be if that variable were uncorrelated with the other variables X_{-k} in the equation. In Table 9, we report the VIFs for the main independent variables in the regressions of Table 8.

Column (1) shows that the VIFs for discrepancy in lot size and years of schooling are 1.14 and 1.60, respectively. Following Baum (2006), a rule of thumb is that if the VIF

[5] Given the small sample size relative to the number of streets (56 clusters), street fixed effects are not controlled for in our regressions. Nevertheless, we control for neighborhood fixed effects (there are 16 neighborhoods).

Table 9
Variance inflation factors for the main independent variables.

	(1) $(m_s^2 - m_a^2)$	(2) $(m_s^2 - m_a^2)/m_a^2$	(3) $\mid m_s^2 - m_a^2 \mid$	(4) $\mid (m_s^2 - m_a^2)/m_a^2 \mid$
Discrepancy in lot size[a]	1.14	1.33	1.13	1.34
Tenure	1.62	1.64	1.62	1.64
Years of schooling	1.60	1.60	1.60	1.60
Property title	1.21	1.21	1.21	1.21
ln(PCE)	1.72	1.72	1.72	1.72
Household head	1.17	1.17	1.16	1.17

Note. Variance inflation factors based on the regressions in Table 8.
[a] The discrepancy in lot size is defined differently in every column.

for a variable is greater than 10, multicollinearity is likely to be a concern. Fortunately, none of the VIFs is higher than 10, and all of them are lower than 1.75.

Overall, the analysis of the VIFs suggests that the result about tenure being the only statistically significant factor in explaining the discrepancy between appraisers' and owner's is not driven by multicollinearity.

3.3. Determinants of housing value

In this section, and following Jimenez (1982), we aim at determining the role of different characteristics of dwellings on their values, and whether the use of owner versus appraiser valuation changes this relationship.

We estimate housing prices equations based on hedonic price analysis (Rosen, 1974), that is, housing prices measures are regressed on housing and neighborhood characteristics. While Jimenez just had one available measure for each housing characteristic, we have two available independent measures, one reported by the owner and the other reported by the appraiser, for several housing characteristics: size of the lot, an indicator whether there is a bathroom within the dwelling, the presence of sewerage system, and indicators of cement walls, cement roof, and cement floor.

Table 10 presents the results of the hedonic price equations. Column (1) shows the results from a hedonic price equation using only information provided by the owner. Our estimates suggest that a 1% increase in the lot size is related to a 0.32% increase in the price of the house. Homes with a sewerage system have, on average, 40% higher value than those without it, while homes with cement roof are worth, on average, 50% more than those that use makeshift roofing. Finally, homes with cement walls have a value difference, on average, of 71% with respect to those that use carton metal sheet walls.

In column (2) we use the same variables as in column (1) but we use the appraiser's home value estimate rather than the owner's estimate. The results appear to be very similar, at least in qualitative terms. However, the owner's estimate of the lot size does not predict the appraiser's home value. This suggests that the owner's estimate of the lot size is very noisy.

Comparing column (3) with column (1) reveals that including the owner's estimate of the size of the lot does not seem to be crucial when predicting housing valuation by the owner. The adjusted-R^2 does not change too much from the model in column (1), which includes the owner's estimated lot size (0.34) to the model in column (3), which excludes it (0.30). If instead of using the characteristics reported by the owner we use the ones reported by the appraiser, a similar picture emerges: the adjusted-R^2 does not change too much from the model in column (4), which includes the appraiser's estimated lot size (0.25), to the model in column (5), which excludes it (0.22). Again, this finding is consistent with our suspicion that the owner's lot-size estimate is a very noise measure of the actual lot size.

Looking at column (6), we observe that 84% of the variation in appraised value can be explained by the appraiser's assessment of lot size, whether the property has an interior bathroom, a sewerage system, cement floors, cement walls and a cement roof. The high explanatory power of the regression reported in column (6) contrasts with the 30% of the variation in self-reported value explained by the same housing characteristics but reported by the owner, column (1). The more than 40% difference in the explanatory power between regressions in column (1) and column (6) is likely to be driven by three main factors: first, while owners are a heterogeneous group, the assessor is just one person. The owner-idiosyncratic component cannot be accounted for in column (1), since we cannot use owner-fixed effects, but the appraisal-idiosyncratic component can be absorbed by the constant term in column (6); second, the real estate agent is likely to follow a highly routinized appraisal procedure; and third, the real estate agent is likely to have both higher ability and better skills in housing valuation (indeed, this is his main task in his job), and he is also likely to have higher information regarding general housing market conditions (remember that our estimate of the intensity of housing transactions is between 2% and 4.75%).

Column (7) reports the estimates from a regression of the home value estimated by the owner on the lot size estimated by the appraisal and on the rest of housing characteristics reported by the owner. The reason for performing such a regression is twofold: first, we have already seen that the appraiser's estimate of the size of the lot is more reliable than the one reported by the owner of the house; second, owners should be more accurate in knowing whether the dwelling has a bathroom inside the house, whether the roof, floor and walls are made of cement,

Table 10
Hedonic price equations. OLS estimates.

Dependent variable	Owner's characteristics			Appraiser's characteristics			Owner's characteristics but appraiser's ln(lot size)	
	(1) $\ln(P_s)$	(2) $\ln(P_a)$	(3) $\ln(P_s)$	(4) $\ln(P_s)$	(5) $\ln(P_s)$	(6) $\ln(P_a)$	(7) $\ln(P_s)$	(8) $\ln(P_a)$
ln(lot size)	0.323***	−0.035	—	0.491***	—	0.637***	0.406***	0.808**
	(0.087)	(0.061)		(0.174)		(0.061)	(0.135)	(0.104)
Bathroom	0.178	0.200**	−0.282*	0.605*	0.612*	0.334***	0.260*	0.145*
	(0.134)	(0.097)	(0.142)	(0.335)	(0.342)	(0.067)	(0.139)	(0.074)
Sewerage	0.396*	0.212**	0.437**	−0.329	−0.369	−0.215*	0.421**	0.176*
	(0.199)	(0.089)	(0.200)	(0.617)	(0.743)	(0.124)	(0.152)	(0.093)
Cement floor	0.067	0.319**	0.165	−0.998***	−**	−	0.138	0.255**
	(0.494)	(0.145)	(0.495)	(0.367)	(0.413)	(0.086)	(0.496)	(0.119)
Cement walls	0.707**	0.378*	0.627*	1.153***	1.302***	0.467	0.536*	0.205
	(0.338)	(0.190)	(0.336)	(0.368)	(0.409)	(0.079)	(0.319)	(0.125)
Cement roof	0.494***	0.554***	0.527***	0.345**	0.439***	0.964	0.491***	0.479***
	(0.142)	(0.096)	(0.149)	(0.139)	(0.129)	(0.063)	(0.149)	(0.075)
Neighborhood dummies?	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
p-Value *F*-test	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Test: equality of coefficients								
			p-Value (1) = (2)				*p*-Value (7) = (8)	
ln(lot size)			0.00					0.02
Bathroom			0.87					0.37
Sewage			0.36					0.22
Cement floor			0.57					0.79
Cement walls			0.26					0.23
Cement roof			0.69					0.94
R^2	0.39	0.42	0.36	0.31	0.28	0.84	0.38	0.57
Adjusted-R^2	0.34	0.36	0.30	0.25	0.22	0.83	0.32	0.53
Sample size	259	259	259	259	259	259	259	259

Note. All regressions include a constant term. Neighborhood dummies: 16 dummies based on the first three digits of the AGEB (Mexican geographical unit) classification system. Clustered robust standard errors are reported in parentheses (52 clusters).
* Significant at the 10%.
** Significant at the 5%.
*** Significant at the 1%.

whether they have a sewerage system, and so forth. Reassuringly, the coefficients in column (7) are very similar to those reported in column (1). For the sake of comparison, column (8) reports a similar regression but using the appraised home value as a dependent variable.

Finally, the table reports the *p*-values for the equality of the coefficients between columns (1) and (2) and between columns (7) and (8) in order to investigate whether the relationships between home value and their characteristics are sensitive to the measure of home valuation. Interestingly, we reject the hypothesis that the relationship between home value and size of the lot is the same depending on the home value measure used in the analysis. However, we cannot reject the hypothesis that the relationships between home value and other housing characteristics are the same depending on the home value measure used in the analysis. These findings are consistent with the view that the real estate agent provides more sensible estimates of the lot size than the owners, and with the fact that he correctly determined housing characteristics without entering the properties.

4. Application: calculating average home values

In this section, mean house values are estimated at the census tract level to study the performance of the self-reported value depending on tenure status. As we argued earlier, self-reported value among homeowners with short tenure provides the preferred estimate of house value. The sample at hand is hardly adequate for the short-tenure measure because we obtained mean home value for nine census tracts from only 12 observations. However, as Table 11 shows, in terms of both bias and inaccuracy, using the short-tenure responses provides superior estimates to using the set of all responses.

The two measures of bias (mean difference and mean percent difference) show that using only the information for short-tenure homeowners provides a less biased estimate of the mean house value. The inaccuracy (measured by both the mean absolute difference and the mean absolute percent difference) is also improved when using only information on home valuation by short-tenure homeowners.

5. Comparing the real estate agent's appraisals to market prices

The analysis presented in Section 3 relied crucially on the assumption that the appraisal's estimated values are unbiased estimates of the market price of the house. Hence, the discrepancy between home values obtained from the owners and the real estate agent is interpreted

Table 11
Estimation of mean house values at the census tract level.

Census tract	Average appraiser's estimate	Average owner's estimate	Average short-tenure owner's estimate
1	9789	26,406	12,636
2	7045	14,072	3636
3	21,221	27,480	37,576
4	8724	21,127	9091
5	10,743	14,321	14,318
6	12,395	15,680	10,152
7	9583	14,613	13,636
8	14,082	14,267	16,667
9	11,074	12,215	3030
Mean difference		6169	1787
Mean % difference		0.63	0.07
Mean absolute difference		6169	4,831
Mean absolute % difference		0.63	0.38

Note. The lower panel means are calculated over the groups in the top panel.

Table 12
Descriptive statistics of different house value estimates.

	Declared transaction price (P_m)	Tax value (P_t)	Appraised value (P_a)	Owner's current valuation (P_s)
Mean	19,759	18,867	35,572	30,202
SD	15,080	11,967	25,561	20,078
Observations	27	27	27	27
Test p-Value	$P_m = P_t$ (0.55) $P_m = P_a$ (0.00) $P_m = P_s$ (0.00)	$P_t = P_a$ (0.00) $P_t = P_s$ (0.00)	$P_a = P_s$ (0.22)	

Note. Test of equality of means using all paired observations available for both measures.

as originating from homeowner's misperceptions about home market value. At the end of Section 2, we offered several reasons for why we think that that is a reasonable assumption. In order to check the validity of this assumption, we sought to compare the agent's appraisals to market prices. The data collection and comparison results are presented in this section.

Property transactions registered in 2006 in Acayucan were obtained from the Public Registry of Property. Each reported transaction had an address, owner's name, sale price, and a city assessed value of the property. Municipal property tax in Acayucan is charged using the maximum of the last reported transaction price and the city assessed value of the property. Only urban houses were considered in the study, not empty lots or rural properties reported in the Public Registry. Inherited or donated homes were ignored, because the focus of the validation was on properties that had been sold.

A professional survey team visited the homes that satisfied these characteristics in early 2007 and administered a short survey. If the owner or his spouse were not available, the survey was not administered. The survey inquired about the date the property was acquired, if any improvements had been made, how much they thought the current value of the property was, and how much they had paid for it. Table 12 shows descriptive statistics on the four available home value measures for the set of homes obtained from the Public Registry. The tests of equality of means reported in the lower part of the Table indicate that the transaction price declared by the homeowners is significantly lower than the appraised value. Assuming that the transaction prices are the best indicator of market prices, the data would lead us to conclude that the appraiser significantly overestimates the value of homes. In light of the results of Section 3, homeowners would seem to have an even larger positive bias than we found earlier.

However, assuming that interviewed homeowners are reporting the price they actually paid may be excessively naïve. A comparison of the first and fourth columns of the Table indicates that: either homes in Acayucan have benefitted from a real home appreciation rates in excess of 26% per year or that homeowners are not reporting the real transaction price in the interview.

During the course of the fieldwork, municipal tax authorities, public notaries, and the real estate agent all mentioned that the tax values obtained from the Public Registry would be substantially lower than the market prices because people tended to underreport the true sale price of the house. If people underreported in order to pay fewer taxes, then it is very likely that they also underreported to the field workers when they were surveyed. The declared market price obtained from the interview is indistinguishable from the tax value of the property. The discrepancy between the tax value and the appraised value is around 44% of the appraised value. It seems then that the obtained "market prices" are suspect, preventing us from credibly using this data source to test whether the real estate agent provides unbiased and accurate estimates of market home values. The third and fourth columns of the Table suggest, as was found in Section 3, that among recent homeowners, appraised values are close to the self-valuations of property value.

Although we made every effort possible to obtain market prices to benchmark the professional appraised values, obtaining market prices in this context proved elusive. This fact underscores the importance of our paper. Homeowner estimated market values seem to be the only widely available home value measure in a developing country such as Mexico, and testing its reliability is important before widely using this measure of house values. Finally, although we could not test whether the assessed home value provides us with an unbiased estimate for the market value, it is important to point out the empirical evidence reported by Clapp and Giaccotto (1992), who compare repeat sales with assessed values, suggests that the effect of measurement error associated with assessed value is negligible.

6. Conclusions

This paper inquired how reliable homeowners' estimates of the value of their houses are in the context of a developing country. Remarkably, the distribution of the bias in owner–appraiser valuations in Acayucan is very similar to the one reported for the Philippines in the study by Jimenez (1982). Hence, our findings on the lack of precision of homeowners' estimates confirm existing results. We also find that the valuation bias associated with longer tenure is positive, confirming the results found in the US studies for recently transacted homes. Our main contribution is the finding that the tenure driven bias is potentially much larger in a developing country context. The results show that, in our sample, owners with long tenure largely overestimate the value of their home, with a mean absolute percent error on the order of 150%. However, families with tenure of less than 2 years have reasonably accurate and unbiased estimates of the value of their home; similar to what is found in the US literature. A cluster of similar homes built by a specialized construction company shows zero bias and dramatically improved precision of estimation with respect to the other subgroups.

We found item non-response to be uncorrelated with the appraised value of the house and other demographic characteristics; this suggests that unbiased estimates of the average value of group of homes can be obtained through household surveys. This is useful because if this were not the case, responses would have to be adjusted using sample selectivity methods which always rely on specific functional form assumptions that cannot be tested and can drive entirely the results.

However useful surveys may be to estimate average home values in a group of homes, though, anyone using the estimated home value for studies of individual behavior should be aware that the mean absolute percentage error was found to be between 50% and 60% of the appraised home value among short-tenured homeowners in this study. Although it is also reassuring that bias and inaccuracy were not robustly related to socioeconomic characteristics, like family income or level of education of the respondent.

To sum up, the results of this study caution against using homeowner estimates for analysis of individual behavior, but suggest that these estimates can be used to reasonably approximate mean home values for groups of homes (for example, at the census tract level). If the objective is to estimate average home value, then the answers from homeowners with short tenure may be used successfully in future work in surveys in developing countries.

Acknowledgments

We thank Orley Ashenfelter, David Atkin, Rodrigo Barros, Angus Deaton, Jaume García Villar, Libertad González, Jon Guryan, Peter Kennedy, Alan Krueger, Adriana Lleras-Muney, Núria Rodríguez, Harvey Rosen, Cecilia Rouse, Michael Rothschild, Jesse Rothstein, all the seminar participants at the Princeton Labor Lunch, Universitat Pompeu Fabra Applied Lunch, Princeton Public Finance Workshop, Lincoln Institute of Land Policy, and the Princeton Research Program in Development Studies Lunch, and several anonymous reviewers for their comments and suggestions. We also specially thank INSAD (Investigación en Salud y Demografía) and his director Jose Luis Palma. Gonzalez-Navarro gratefully acknowledges financial support from the Fellowship of the Woodrow Wilson Scholars (2006–2008) and Quintana-Domeque from the Bank of Spain (2006–2007), and the Rafael del Pino Foundation (2007–2008). This project is funded by the Princeton Center for Health and Wellbeing, Princeton Industrial Relations Section, Princeton Research Program in Development Studies and the Lincoln Institute of Land Policy. The data collected for this study underwent the approval process of Princeton University's Institutional Review Panel (Research Protocol 3104). A working paper version of this article was previously published in the Institut Valencià d'Investigacions Econòmiques (IVIE) AD Series: IVIE Working Paper WP-AD 2008-18. All remaining errors are ours.

References

Agarwal, S., 2007. The impact of homeowners' housing wealth misestimation on consumption and saving decisions. Real Estate Economics 35 (2), 135–154.

Baum, C., 2006. An Introduction to Modern Econometrics Using Stata. Stata Press.

Campbell, J.Y., Cocco, J.F., 2007. How do house prices affect consumption? Evidence from micro data. Journal of Monetary Economics 54 (3), 591–621.

Clapp, J.M., Giaccotto, C., 1992. Estimating price indices for residential property: a comparison of repeat sales and assessed value methods. Journal of the American Statistical Association 87 (418), 300–306.

David, E.L., 1968. The use of assessed data to approximate sales values of recreational property. Land Economics 44 (1), 127–129.

DiPasquale, D., Somerville, C.T., 1995. Do house price indices based on transacting units represent the entire stock? Evidence from the American Housing Survey. Journal of Housing Economics 4 (3), 195–229.

Disney, R., Bridges, S., Gathergood, J., 2006. Housing Wealth and Household Indebtness: Is there a Household 'Financial Accelerator'? Unpublished manuscript. University of Nottingham, Nottingham.

Gonzalez-Navarro, M., Quintana-Domeque, C., 2007. Description of the Acayucan Standards of Living Survey. Unpublished manuscript. Princeton University, Princeton, NJ.

Goodman, J.L., Ittner, J.B., 1992. The accuracy of home owners' estimates of house value. Journal of Housing Economics 2 (4), 339–357.

Jimenez, E., 1982. The value of squatter dwellings in developing countries. Economic Development and Cultural Change 30 (4), 739–752.

Juster, F.T., Lupton, J.P., Smith, J.P., Stafford, F., 2005. The decline in household saving and the wealth effect. Review of Economics and Statistics 87 (4), 20–27.

Kain, J.F., Quigley, J.M., 1972. Note on owner's estimate of housing value. Journal of the American Statistical Association 67 (340), 803–806.

Kiel, K.A., Zabel, J.E., 1999. The accuracy of owner-provided house values: the 1978–1991 American Housing Survey. Real Estate Economics 27 (2), 263–298.

Kish, L., Lansing, J.B., 1954. Response errors in estimating the value of homes. Journal of the American Statistical Association 49 (267), 520–538.

Klyuev, V., Mills, P., 2006. Is housing wealth an 'ATM'? The relationship between household wealth, home equity withdrawal and savings rates. International Monetary Fund Working Paper WP/06/162, June.

Lusardi, A., Mitchell, O.S., 2007. Baby boomers retirement security: the role of planning, financial literacy and housing wealth. Journal of Monetary Economics 54 (1), 205–224.

Martinelli, C., Parker, S.W., 2009. Deception and misreporting in a social program. Journal of the European Economic Association 7 (4), 886–908.

Rosen, S., 1974. Hedonic prices and implicit markets. Journal of Political Economy 82 (1), 34–55.