Labor Income, Housing Prices, and Homeownership

Thomas Davidoff

Haas School of Business, UC Berkeley *

March 18, 2005

*Mailing address: Haas School of Business, UC Berkeley, Berkeley, CA 94720. This paper is an update of the first chapter of my doctoral thesis at MIT. I am grateful to Jan Brueckner, Peter Diamond, Sendhil Mullainathan, Jim Poterba, William Wheaton and participants at seminars at MIT, Berkeley, Columbia, Wharton, the Federal Reserve Board, Syracuse University and George Washington University for helpful suggestions.

Abstract

This paper tests the intuition that households whose incomes covary relatively strongly with housing prices should own relatively little housing. Among US households, a one standard deviation in covariance between income and home prices is associated with a decrease of approximately $7,500 in the value of owner occupied housing. This result arises in the presence of controls for the level and distribution of home prices. The generally positive correlations between income and home prices suggests that households enter financial markets with a greater exposure to risk than is typically modeled.

1 Introduction

This paper asks whether households whose incomes covary relatively strongly with housing prices purchase relatively little housing. The question is motivated by the fact that for many households, the home is the largest asset in the portfolio and labor income is the most important source of wealth.[1] Without financial markets for insurance against volatility in housing prices and labor income, such as those proposed by Shiller [19] and Caplin et al. [5], it is natural to expect risk averse households to use housing purchases to hedge income risk.

Households can "purchase relatively little housing" either by deciding to rent rather than own (acting on the "extensive margin") or by purchasing a comparatively inexpensive home (acting on the "intensive margin"). I estimate the effect of income-price covariance on both margins.

Existing empirical tests of whether people diversify away labor income risk focus on stock market behavior, and ignore housing as a form of investment.[2] In light of the dominant share of housing relative to stocks in most portfolios, this paper takes the opposite approach and assumes for simplicity that workers save or borrow only through housing purchases and riskless assets. The dominance of housing is established by Aizcorbe et al. [1], who use the 1998 Survey of Consumer Finances to estimate that 66 percent of households owned their own home in 1997, but only 56 percent did any saving and just 49 percent held any stock, directly or through mutual funds or retirement plans. Among homeowners, the median home value was $100,000, whereas the median value of equities among shareholders was $25,000. In the 1990 US Census data considered here and tabulated below, median 1989 investment income was just $10 for homeowners and zero for renters.

A growing body of theoretical and empirical work considers the role of housing in the portfolio. An early observation, developed by Henderson and Ioannides [14], is that some

[1] See, e.g. Aizcorbe et al. [1].

[2] Heaton and Lucas [13] find in a panel of US investors, the fraction of wealth put into stocks decreases in the covariance between entrepreneurial income and stock market returns. By contrast, Vissing-Jorgenson [23] fails to find such an effect.

homeowners will own too much housing because their consumption demand for housing exceeds their investment demand. This possibility arises under the realistic assumption that homeowners are unable to sell off any of their home equity, except by becoming renters and owning zero housing.[3] Brueckner [3] extends this analysis to consider the case in which consumers hold a composite portfolio of housing, stocks and riskless bonds. Brueckner's result is that homeowners' portfolios may be mean-variance inefficient in the sense that a sale of housing in exchange for some set of stocks would increase both expected return and reduce variance. An inefficient portfolio is held if demand for housing consumption exceeds the mean-variance efficient quantity of housing. A similar result is shown in simulations presented in Flavin and Yamashita [10].

Several empirical papers lend support to the idea that housing considerations affect portfolio choice. Fratantoni [11] and Yamashita [24] find that housing crowds out investment in stocks. Chetty and Szeidl [8] find that more recent home buyers hold fewer stocks in their portfolios than homeowners with longer tenures and claim that this is a result of increased risk aversion generated by large mortgage debt. Sinai and Souleles [20] find that variance in housing prices is associated with increased homeownership and higher housing prices, and attribute this to increased demand for hedging rental expenditure risk.

Some recent theoretical papers consider housing choice in the context of both uncertain housing prices and uncertain labor income. Lustig and Nieuwerburgh [15] and Piazzesi et al. [18] discuss the macro consequences of housing risk. Campbell and Cocco [4], Cocco [9] and Yao and Zhang [25] solve numerically for optimal lifetime mortgage and housing behavior. These papers estimate a single population covariance matrix for prices, labor income and interest rates (and zero-covariance stocks in the case of Yao and Zhang) and assume jointly normal distributions. By contrast, I confine the theoretical analysis to a two period setting, allow for population heterogeneity in the covariance between labor income and housing prices and describe analytically conditions under which housing purchases fall with covariance. In a special case, Ortalo-Magné and Rady [17] show that renting becomes becomes relatively

[3] Only a trivial number of renters own rental housing.

more attractive than homeownership as the covariance between income and housing prices increases.

Following this introduction, the second section of this paper presents the theoretical model and shows that consumers with mean-variance preferences optimally purchase less housing as the covariance between labor income and housing prices rises. I also describe more general conditions under which the extensive margin result of Ortalo-Magné and Rady [17] holds. In the third section, I describe the data used to test these theoretical results. By estimating a separate income-price covariance for each industry in each US metropolitan area (MSA), I am able to exploit variation across both industries and regions. Thus, the estimates of the effect of income-price covariance on housing purchases are conditional on both the level and variance of housing prices because fixed effects for each MSA are present. Industry fixed effects allow identification even if the types of people who work in industries that typically have wage movements highly correlated with housing prices have different housing demands than other consumers. The fourth section details the results, which confirm the theoretical predictions. Notably, a one standard deviation increase in income-price covariance is associated with a reduction in the value of housing owned of approximately $7,500. The fifth section concludes.

2 Housing Choice with Stochastic Labor Income and Prices

2.1 Model Set-Up and Assumptions

Present housing decisions affect lifetime utility directly through the benefits of consuming more or less housing and indirectly through the lifetime budget constraint. With complete markets, these effects would be separable and consumers would be able to own any desired fraction of the value of the home in which they lived. Realistically, most homeowners, for whatever institutional or behavioral reasons, are not able to separate housing investment and consumption.[4] I strengthen the assumptions of Henderson and Ioannides [14] so that renters all own zero housing and homeowners own exactly as much housing as they consume. In translating the theory to empirical work below, the implicit assumption is that the owners of rental housing are all professional landlords.

Housing choice is considered in a two period world. Consumers derive utility from two goods, housing and a numeraire, in each period. Labor supply is fixed. In the first period, consumers earn labor income y_1, choose a level of savings or debt M, and choose a quantity of housing to consume H. If the consumer rents, she pays $H\mathrm{Rent}_1$ for her housing consumption. Homeowners pay HP_1. Consumers can borrow and lend unlimited amounts at an interest rate of r. The choices of housing consumption and debt or savings imply that first period numeraire consumption is given by:

$$c_1 = \begin{cases} y_1 - HP_1 + M & \text{If homeowner} \\ y_1 - H\mathrm{Rent}_1 + M & \text{If renter} \end{cases} \tag{1}$$

In the second period, all consumers earn y_2 in labor income and repay or earn $(1+r)M$ in principal and interest on debt or savings. Homeowners earn an additional HP_2 in proceeds from sale of their home. y_2, P_2 and r are measured in units of second period numeraire consumption. y_2 and P_2 are stochastic and r is considered riskless both to the borrower and to the lender, so that default is not possible. Second period wealth is divided between numeraire and housing consumption.

We analyze how the covariance $COV(P, y)$ between y_2 and P_2 affects (1) the decision to own or rent housing, and (2) conditional on deciding to own, how much housing to purchase.

This simple two period model involves some simplifications. First, the date at which a home is sold is deterministic and set equal to the length of stay in rental housing. Second period prices can be thought of as an average of prices at all feasible resale dates multiplied by the probability of resale at that date, but we are ruling out use of the move date as a partial hedge against changes in housing prices.

[4] These institutional constraints are discussed at length in Caplin et al. [5].

Housing demand after resale is left unspecified. It will prove analytically convenient to assume that second period housing needs are fixed, particularly at a level of zero.[5] If the consumer has no bequest motive, then renting in the second period will be preferred to owning. We can assume that these rents would be proportional to P_2, justified by the partial equilibrium approach and the presence of regional fixed effects in the empirical work. With a bequest motive, it is possible that purchase would be optimal.

The portfolio choice problem is simplified in that savings are assumed riskless and freely chosen. Risklessness is justified by the fact, noted above, that equities are a small part of most portfolios relative to housing. For some consumers, unlimited borrowing capacity is a poor assumption, so housing purchases may be corner solutions. Hence an empirical relationship between housing choice and income-price covariance might simply reflect some relationship between borrowing constraints and covariance. While it is not clear why such a relationship would exist, the empirical analysis deals with this potential problem by restricting the sample to households unlikely to face a borrowing constraint in some specifications. One might imagine that lenders would impose borrowing constraints on individuals whose incomes move strongly with housing prices, since this would render default more likely.[6] This appears not to be the case, however, since the major US mortgage guarantors, Fannie Mae and Freddie Mac do not question the variance or covariance properties of borrower incomes.

The partial equilibrium approach ignores the possibility that the level of housing prices P_1 as well as the mean and variance of P_2 are affected by the distribution of $COV(P, y)$ across households. We might expect, for example, that in markets where the average covariance is high, housing prices would be low. We might also expect that rents would be relatively high in such markets. This possibility must be dealt with to interpret a negative empirical relationship between $COV(P, y)$ and the dollar value of housing purchases as reflecting the

[5] There is no consensus among housing economists as to how the utility of older homeowners is shaped by housing decisions (see, for example, Venti and Wise [22]), so any life cycle model of housing choice necessarily involves strong assumptions.

[6] Default is more likely with a higher income-price covariance because high covariance borrowers are likely to be underwater in the very states of nature in which their incomes fail to cover mortgage payments.

Table 1: Notation

Variable	Description
H	Housing consumed in the first period (purchased or rented).
M	Mortgage debt (savings if negative).
$u(c_1, H; Z)$	Concave utility function.
y_t	Labor income in period t.
P_t	Price of housing in period t.
Z	Vector of individual characteristics.
v	Indirect utility function, concave in W_2.
$W_2\|_{own} = y_2 + HP_2 - (1+r)M$	Second period wealth for homeowners.
$W_2\|_{rent} = y_2 - (1+r)M$	Second period wealth for renters.
P_2	Second period housing price.
θ	Joint distribution of P_2 and y_2.
r	mortgage interest rate.
$COV(P, y)$	Covariance between second period income and housing price.

mechanism discussed below. This is accomplished by considering cross sectional variation only within markets.

The existence of a single housing price in each market assumes that structure, lot size and other locational characteristics can be aggregated meaningfully.

The notation relating to consumer choice is summarized in Table 1.

2.2 Homeowner Utility Maximization

If a consumer chooses to purchase housing, the quantity H and mortgage M are chosen to maximize expected utility given by:

$$U(H, M|\Theta, Z) = u(y_1 + M - HP_1, H, Z) + Ev(W_2|_{own}, P_2|Z, \Theta). \tag{2}$$

The first order conditions are:

$$U_H = -P_1 u_1 + u_2 + E(P_2 v_1) = 0, \tag{3}$$

$$U_M = u_1 - (1+r)Ev_1 = 0. \tag{4}$$

2.2.1 Effect of increasing covariance on housing purchases

Expected second period utility will, in general, depend on all the moments of the joint distribution of future housing prices and income. If we consider a change in a particular parameter of the joint distribution θ, holding consumer characteristics Z and the rest of the moments Θ constant, then we can think of the other moments as fixed parameters of the utility function. We can thus rewrite expected utility conditional on Z and all of Θ except for θ as

$$U(H, M, \theta).$$

Total differentiation of the first order optimality conditions (3) and (4) gives us two equations in two unknowns, which can be solved jointly for the change in optimal housing purchases associated with a small increase in the parameter θ. These total derivatives are given by:

$$0 = U_{M\theta} + U_{MM}\frac{dM}{d\theta} + U_{MH}\frac{dH}{d\theta}, \tag{5}$$

$$0 = U_{H\theta} + U_{HM}\frac{dM}{d\theta} + U_{HH}\frac{dH}{d\theta}. \tag{6}$$

Combining and rearranging conditions (5) and (6) gives the result:

$$\frac{dH}{d\theta}(U_{MM}U_{HH} - U_{MH}^2) = -U_{H\theta}U_{MM} + U_{MH}U_{M\theta} \tag{7}$$

The term multiplying the derivative of interest $\frac{dH}{d\theta}$ must be positive by concavity of u and v (see, for example, Mas-Collel et al. [16], Appendix D). The second derivative U_{MM} similarly must be negative, so dividing equation (7) by $-U_{MM}$ we have the relation:

$$\text{sign}(\frac{dH}{d\theta}) = \text{sign}(U_{H\theta} - \frac{U_{MH}}{U_{MM}}U_{M\theta}). \tag{8}$$

Intuitively, a parameter shift tends to reduce the quantity of housing if the shift reduces the marginal benefit of housing purchases. This effect is modified by changes in mortgage debt if changes in housing investment affect the marginal benefit of mortgage debt. An induced increase (decrease) in the marginal benefit of mortgage debt tends to increase (decrease) housing purchases if increased housing investment makes mortgage debt relatively attractive. The opposite implications arise if mortgage debt becomes less attractive with housing purchase. In our case, the distributional parameter of interest θ is the covariance between income and prices, $COV(P, y)$.

Sufficient conditions for housing purchases to decrease in covariance exist under a pair of assumptions shared by Berkovec and Fullerton [2] and Flavin and Yamashita [10]. These papers specialize the homeowners' maximization problem by assuming first that there is zero demand for housing in the second period, so that expected indirect utility Ev in equation (2) depends only on the distribution of future wealth. The second assumption is that expected indirect utility depends only and additively on the mean and variance of second period wealth:

$$Ev = a(EW_2) + b(VAR(W_2)); \tag{9}$$
$$a' > 0, b' < 0.$$

With wealth given by $y_2 + HP_2 - (1 + r)M$, and the borrowing rate r fixed between purchase and sale of housing, the variance of future wealth is given by:

$$VAR(W_2) = VAR(y_2) + 2HCOV(P, y) + H^2 VAR(P_2). \tag{10}$$

In this case, an increase in covariance (holding expected income and prices constant) has no direct effect on first period utility or on the value of expected second period wealth. This implies that mortgage debt and covariance do not interact in expected utility. The term $U_{MCOV(P,y)}$ thus equals zero and equation (8) reduces to the effect of covariance on the marginal utility of housing purchases. This effect is given by:

$$\text{sign}(\frac{dH}{dCOV(P, y)}) = \text{sign}(b' \frac{\partial^2 VAR(W_2)}{\partial H \partial COV(P, y)}) = \text{sign}(2b') < 0.$$

Hence, in this setting, optimal housing purchases conditional on owning are decreasing in covariance, matching intuition. We can also see that for constant variance and mean growth in income and prices, for any positive level of housing, the variance of wealth is increasing in the covariance term. Thus, expected utility falls with the covariance for any level of housing. By implication, expected utility conditional on owning must fall.

Both mean-variance utility and the absence of housing consumption after resale are highly restrictive assumptions. Nevertheless, to the extent that we believe homeowners are in a long position in housing and that mean-variance utility is a decent approximation, there is formal justification for taking the intuition to data.

2.3 Renters' Expected Utility

Renters avoid exposure to housing prices in nominal wealth, but face greater expenditure risk to the extent that they continue to consume housing in a market with correlated housing prices. Assume that the first period renters also rent in the second period and that the ratio of rents to prices is g_1 in the first period and a deterministic constant g_2 in the second period. Individual expected utility, having decided to rent in the first period is:

$$EU = \max_{H,M} u(y_1 - Hg_1P_1, H; Z) + Ev(y_2, P_2; Z, \Theta). \tag{11}$$

For renters, second period numeraire wealth and its variance are given by:

$$W_2|_{\text{rent}} = y_2 - H_2 g_2 P_2 - (1+r)M, \tag{12}$$

$$VAR(W_2) = VAR(y_2) + g_2^2 VAR(H_2 P_2), \tag{13}$$

where H_2 denotes second period housing consumption.

If we assume that second period housing needs for renters are fixed at some level $\bar{H}$, then (13) implies:

$$VAR(W_2) = VAR(y_2) + \bar{H}^2 g_2^2 VAR(P_2) - 2\bar{H} g_2 COV(P, y). \tag{14}$$

With the further assumption that renters have mean variance utility over second period numeraire wealth, we obtain the result that renters' expected utility increases with $COV(P, y)$.

The assumption of fixed housing needs may be justified by the stylized fact that there is less variability in the quality of rental housing than in the quality of owner occupied housing. The result of decreasing variance of numeraire consumption with increasing covariance holds more generally if the price elasticity of demand for housing is less than one.

Summarizing, we have the following:

Result 1 *If second period housing consumption is fixed and if indirect utility over second period numeraire consumption is additively separable mean-variance, then*

a. Intensive Margin: *Optimal first period housing purchases conditional on purchase are decreasing in $COV(P, y)$.*

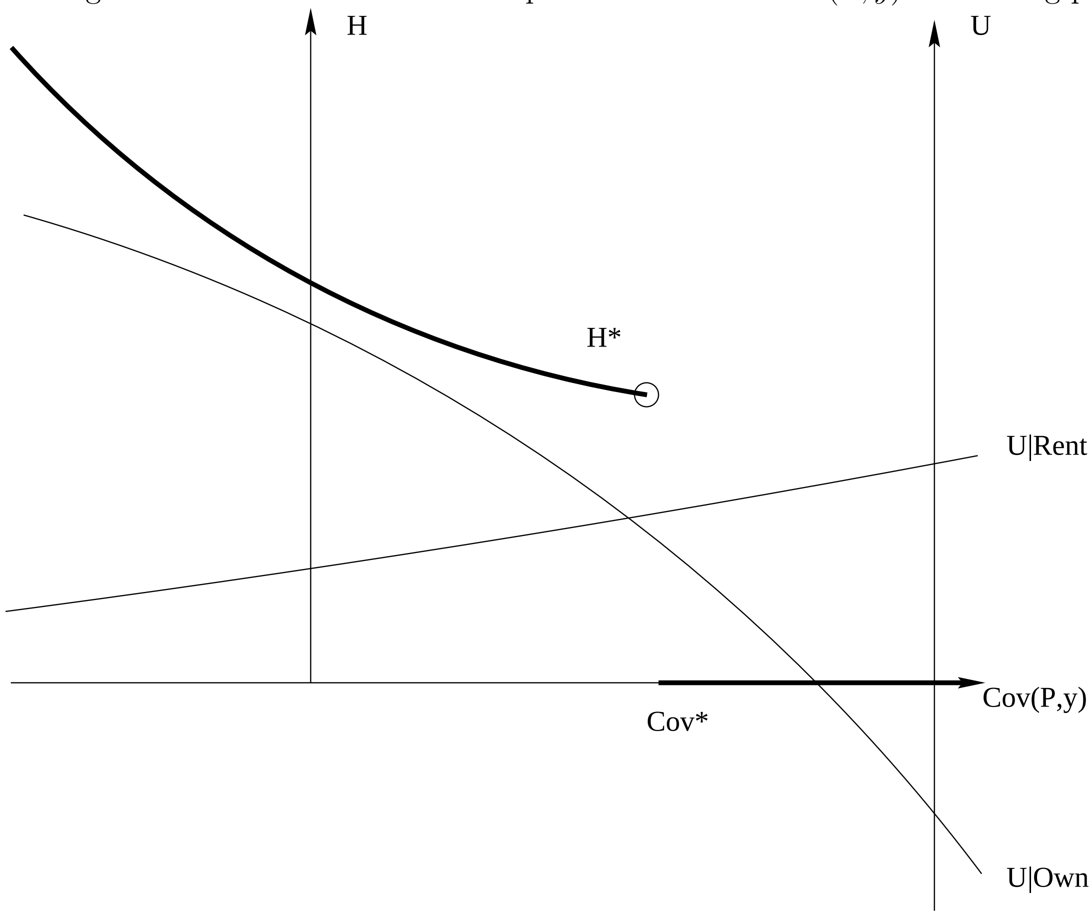
b. Extensive Margin: *The difference in maximized utility conditional on owning and maximized utility conditional on renting falls with $COV(P, y)$. Hence, with sufficient variation in $COV(P, y)$, conditional on characteristics there is a critical value for $COV(P, y)$, $Cov*$ above which renting is optimal and below which owning is optimal.*

Result 1 can be portrayed graphically as in Figure 1. H^* represents optimal first period housing purchases and $U|Rent$ and $U|Own$ are maximized utility conditional on owning or renting. The thick black line represents optimal housing purchases; renting housing is deemed a zero purchase of housing. The curvatures of conditional utility and housing purchases are based on speculation. Nonlinear effects of income-price covariance on housing purchases and on the decision to own or rent are considered empirically below.

3 Empirical Approach and Data

I test Result 1 in several ways. Following the solid line $H*$ in Figure 1, the theory suggests that within a housing market, the value of housing owned (zero for renters) should decrease

Figure 1: Effect of labor income - price covariance $COV(P, y)$ on housing purchases



in $COV(P, y)$. I thus present regressions of the form:

$$VALUE = b_0 + b_1 COV(P, y) + Zb_2 + \epsilon, \tag{15}$$

where Z is a vector of covariates. The theory is consistent only with a negative value for b_1. ϵ captures idiosyncratic differences in demand for owner occupied housing. Because the value of purchased housing does not go smoothly to zero in the population, we know that there will not be a linear effect of $COV(P, y)$, so estimation of equations of the form (15) must allow for heteroskedasticity.

A second way to test the theory is to test directly the effect of housing purchases on the decision to own or rent. Denoting the choice to rent, rather than own, by $RENTER$ and the standard normal cumulative distribution function by Φ, I thus present estimates of the form:

$$Pr(RENTER) = \Phi(\delta_1 COV(P, y) + \delta_2 Z). \tag{16}$$

Here, we expect a positive sign on δ_1.

The theory predicts that increasing income-price covariance will lead to decreased purchases of housing conditional on deciding to own and conditional on all characteristics. This leads to regressions restricted to homeowners of the form:

$$VALUE|OWN = \beta_0 + \beta_1 COV(P, y) + Z\beta_2 + \epsilon, \tag{17}$$

Such a regression runs into the following problem of selection on unobservables. Individuals who decide to own housing despite a large covariance between price and income presumably do so because they have a taste for owner occupied housing (a large value of ϵ). Hence as $COV(P, y)$ grows, the average value of ϵ grows among the set of owners. This should bias the estimated value of β_1 towards zero, and away from the theoretical prediction of a negative effect. A significant negative value of β_1 is thus confirmation of the theory. I also present a set of Heckman sample selection correction estimates, thereby incorporating deviations from predicted homeownership from equation (16) into an additional regressor in equation (17).

3.1 Data

Equations (15), (16) and (17) are estimated using three sources of data. The dollar value of housing owned and owner or renter status are identified from the University of Minnesota's Integrated Public Use Microdata (IPUMS) 1990 Census data, a sample of approximately five percent of US households. The IPUMS data also indicates the industry (2 digit SIC) in which respondents work, the metropolitan area (MSA) in which they live, as well as a large number of demographic characteristics, summarized in Table 4. I confine the IPUMS data to household heads who report the industry (two digit SIC) in which they work, are less than 62 years old and who live in an MSA with house price data. This leaves approximately

1.1 million household heads. Table 3 highlights the importance of housing as an asset in the data.

MSA and SIC fixed effects are included in the demographic variables (Z in equations (17) through (15)). Where appropriate, all demographic variables are included along with their interaction with income. Interacting income with MSA fixed effects allows for different price effects on housing owned at different income levels.

The correlation and covariance between income and housing prices, as well as the variance of income are estimated by merging two additional data sets. A panel of mean wages by industry and MSA (ES202) from 1976 to 1999 was obtained from the Bureau of Labor Statistics. Comovement between wages and MSA housing prices is estimated by merging the ES 202 data with the Office of Federal Housing Agency Oversight (OFHEO) repeat sale house price index (HPI). The structure of the data allow a different covariance or correlation to be estimated for each MSA-SIC cell. So, for example, the estimate of the correlation between prices and incomes for retail workers in Boston is different from the estimate of correlation between prices and incomes for construction workers in Boston and different from the estimate of correlation for retail workers in Detroit.

To estimate the effect of $COV(P, y)$ on housing choice, the IPUMS data is merged with the ES 202 - HPI covariance data in the way discussed below. There are 148 MSAs with HPI data and 68 SIC codes with ES 202 data. Some MSA-SIC cells are deleted for lack of sufficient observations (less than 200 members in the IPUMS data), so regressions are based on individuals in approximately 7,400 cells. Income and price data are deflated by the US consumer price index for non-housing goods.

3.2 Estimation of income-price covariance terms

Income is observed only for 1989 in the IPUMS data. For other years t, I assume that income for individuual i is given by:

$$y_{it} = y_{IPUMSi} \frac{y_{ES202t}}{y_{ES202,1989}} + \epsilon_{it}, \tag{18}$$

where y_{ES202t} is the mean income in year t for the MSA-SIC cell in which individual i works. I assume that the idiosyncratic portion of individual incomes ϵ_{it} is not correlated with local housing prices. Relaxation of this assumption is discussed below.

The correlation between income and price ($CORR(P, y)$) is measured as follows. A series of percentage changes in wages and house price index values is created from the ES 202 and HPI data and the correlation between percentage changes in wages and percentage changes in prices is estimated separately for each MSA-SIC cell. This correlation is attributed to all workers in a given MSA-SIC cell. Mathematically, the calculation of correlation follows the following equations:[7]

$$GROW(P) = \sum_{t=1981}^{1999} \frac{HPI_t}{HPI_{t-5}}/18 \tag{19}$$

$$GROW(y_{ES202}) = \sum_{t=1981}^{1999} \frac{y_{ES202t}}{y_{ES202t-5}}/18 \tag{20}$$

$$VAR(P) = \sum_{t=1981}^{1999} (\frac{HPI_t}{HPI_{t-5}} - GROW(P))^2/18 \tag{21}$$

$$VAR(y_{ES202}) = \sum_{t=1981}^{1999} (\frac{y_{ES202t}}{y_{ES202t-5}} - GROW(y_{ES202}))^2/18 \tag{22}$$

$$CORR(P, y_{ES202}) = \frac{\sum_{t=1981}^{1999}(\frac{y_{ES202t}}{y_{ES202t-5}} - GROW(y_{ES202}))(\frac{HPI_t}{HPI_{t-5}} - GROW(P))}{\sqrt{VAR(P)}\sqrt{VAR(y_{ES202})}} \tag{23}$$

The cell-level covariance between income and prices is given by:

$$Cov(P, y_{ES202}) = CORR(P, y_{ES202})\sqrt{VAR(P)}\sqrt{VAR(y_{ES202})}. \tag{24}$$

Whereas $CORR(P, y_{ES202})$ and $COV(P, y_{ES202})$ are constant across individuals within an MSA-SIC cell, individual level $COV(P, y)$ is not. This is because the variance of wages is greater for individuals with higher income than for those with lower income. $COV(P, y)$ is obtained by multiplying $CORR(P, y_{ES202})$ by the standard deviation of MSA price growth

[7]When there are missing years of income data, the means are revised accordingly. Because of the frequently small number of observations of y_{ES202}, variance terms are not multiplied by $\frac{n}{n-1}$.

from the HPI data, by the standard deviation of cell mean wage growth from ES 202 data and by reported IPUMS income. This is necessary because y_{ES202} is an index common to all workers in an MSA-SIC cell, but the wage level varies across workers in the IPUMS data. Expanding (24) and using obvious notation:

$$COV(P, y) = CORR(P, y_{ES202}) \times \sqrt{VAR(P)} \times \sqrt{VAR(y_{ES202})} \times y_{IPUMS}. \tag{25}$$

$VAR(P)$ and $VAR(y_{ES202})$ are estimated using the same data series as $CORR(P, y)$. $VAR(P)$ is the mean squared deviation of MSA percentage house price changes from the MSA's series mean, and hence varies across MSAs but is constant within MSAs. $VAR(y_{ES202})$ varies even within MSAs, but is constant within MSA-SIC cells. y_{ES202} varies within MSA-SIC cells.

A question arises as to which horizon should be used as the basis for variance terms, since wage and price data are available at the quarterly frequency. The estimates are based on overlapping five year horizons for two reasons. First, both data sets are noisy; in general, Griliches and Hausman [12] note that with noisy data, differences are likely better measured over long horizons and Case and Shiller [7] make this point with respect to repeat sale price indices in particular. Second, for homeowners, longer horizons are more relevant since sale within a year or less is quite unlikely. Empirically, covariances over different horizons are highly correlated with each other, so the choice of horizon does not affect the regression results.

The fact that $COV(P, y)$ is defined to be different across individuals within the same MSA-SIC cell is appropriate for regressions with dollar value of housing owned as a dependent variable because the change in the dollar value of housing owned with increasing covariance should be greater for individuals with greater incomes (and greater housing demands). When status as a renter is the dependent variable, it is not clear that the interaction with income is appropriate (if anything, we might expect a weaker effect of covariance for higher income households who are particularly likely to purchase housing), so $COV(P, y)$ is replaced with $COV(P, y_{ES202})$ on the right hand side of the probit regressions of the form (16) and in the

Heckman sample selection regressions.

3.3 Identification and Inference

This section discusses two challenges to interpreting the results of estimating equations (17) through (15). The first challenge is to establish that a negative relationship between income-price covariance and housing purchases is an empirical confirmation of the theory. The second challenge is to interpret the coefficient on $COV(P, y)$ given measurement error in both individual income and in the covariance between MSA-SIC cell income and MSA price as well as conceptual measurement error related to the formulation of $COV(P, y)$.

The results of estimating equations (17) through (15) can only be interpreted as relating to the theoretical predictions if we can rule out non-portfolio reasons for a negative relationship between housing purchases and income-price covariance. The inclusion of demographic covariates in the regressions is important in this way. As noted in the introduction, there is no evidence to suggest that lenders constrain individuals with high income-price covariance to borrow less money. To check for the possibility that liquidity constraints have an empirical correlation with income-price covariance that drives the results, I perform a robustness check by repeating the regressions with the sample confined to workers older than 45. Older workers presumably do not face lifecycle-based liquidity constraints.

An assumption of identical housing price levels (normalized to one) across MSAs is implicit in equation (25). Given the differences in amenity across MSAs, it would be difficult to estimate different hedonic prices across MSAs. Inclusion of MSA fixed effects and interactions with income preserves identification of the effects of covariance, even if the MSA price level (however defined) is correlated with income-price covariance.

Equation (25) contains the assumption that there is a single housing market in each MSA. If there are multiple markets within metropolitan areas, then an identification problem could arise if workers in high income-price covariance industries tend to live in low price markets for non-portfolio reasons. Controls for income, education and national industry fixed effects should alleviate any such concerns.

Measurement error in census reported income and in income-price covariance can be expected to bias the effect of income-price covariance to zero. Income in equation (25) should be equal to a lifetime average income. Reported census income deviates from lifetime income both because reported income in 1989 may deviate from true income in 1989 and because 1989 income is not the same as lifetime income. Based on the estimates of Solon [21], if $Cov(P, y_{ES202})$ were perfectly measured, we would thus expect the coefficient on the interaction to be reduced by up to 50 percent.

To illustrate and address the measurement error in $COV(P, y_{ES202})$, I create ten instruments for each observation of this variable. The instruments are created as follows: for each MSA-SIC cell, I identify the ten geographically closest MSAs that have both house price data and wage data for the same SIC group, based on Cartesian distances between MSA centroids. The estimated correlations $CORR(P, y_{ES202})$ for each of these ten nearby MSA-SIC cells form the instruments. For example, the correlation between mean income and house prices for retail workers in Hartford is one of ten instruments for the covariance between mean income and house prices for retail workers in Boston. Likewise the value of $CORR(P, y_{ES202})$ for Boston retail workers is an instrument for $COV(P, y_{ES202})$ among Hartford retail workers. However, neither the Hartford retail nor the Boston retail value for $CORR(P, y_{ES202})$ is an instrument for the covariance between prices and wages for retail workers in far away Los Angeles. Because the individual level covariance $COV(P, y)$ interacts y_{IPUMS} with $COV(P, y_{ES202})$, the ten correlation instruments are interacted with y_{IPUMS} to form instruments for $COV(P, y)$. For this reason, the instruments for $COV(P, y)$ vary even within MSA-SIC cells.

That $COV(P, y_{ES202})$ is imperfectly measured can be seen in the following exercise. If this covariance measure for a particular MSA-SIC cell is regressed on the covariance for the cell in the same SIC code in the geographically closest MSA, a coefficient of .25 is attained. If the nearest neighboring cell in the same SIC group is instrumented with the second through tenth nearest MSA-SIC cells, a coefficient of 1 is estimated. This can be shown to imply that the noise-to-signal ratio in $COV(P, y_{ES202})$ is close to four-to-one. This, in turn, implies that the estimated coefficient on $COV(P, y)$ in regressions of the form (27) would be just $\frac{1}{4}$ of the true value if there were no other identification problems and if income were perfectly measured.

Some measurement error in income most likely survives the IV strategy and biases down the estimated coefficient on $COV(P, y)$. To check whether measurement error in income somehow drives the results, I also present regressions with no income-$COV(P, y_{ES202})$ interaction that consider the effect of $COV(P, y_{ES202})$ on housing demand, limiting the sample to narrow income bands.

Some conceptual measurement error in $COV(P, y)$ is unavoidable. Different workers in the same region and industry will have different income trajectories because of differences in occupation and experience. Further, workers do not necessarily remain in the same industry or metropolitan area forever. Hence $COV(P, y)$ may either overstate or understate the covariance between an individual's income and housing prices. These facts should not undermine the interpretation of the results, since wages from present earnings should be correlated with lifetime income. We can interpret the regression results as estimating the relationship between housing purchases and the covariance between wages in one's current industry and housing prices in one's current MSA.[8]

In general covariance, rather than correlation is used as a measure of comovement because for a given correlation between prices and wages, increased variances will increase the effect of housing purchases on the variance of second period income. This strategy requires controls for variances and their interactions with income, since the theory shows an effect of $COV(P, y)$ only conditional on variances. The variance of prices and and its interaction are subsumed by MSA fixed effects and their interaction with income. Controls for income variance and its interaction with income are directly included in the estimates presented

[8]It is not clear that it would be preferable to form covariance estimates based on individual income histories, even if such data were available; it is not clear why individual incomes would move with regional home prices except through shared industry or occupation shocks.

below. The use of incomes interacted with neighboring correlations as an instrument should alleviate any concerns that something other than comovement drives the results.

3.3.1 Variance-Covariance Results

Aggregate variance and covariance statistics are reported in Table 4. These statistics arise from a merge of the variance-covariance estimates with income, industry and MSA data from the 1990 US Census five percent state sample from Minnesota's IPUMS database. The extreme correlations of 1 and -1 are observed in a handful of MSA-SIC cells with only two observations of five year changes in wages. Their exclusion does not affect the regression results.

The mean variance of cell income growth ($VAR(y_{ES202})$) is approximately 1 percent, relative to mean growth ($GROW(y_{ES202})$) of approximately 4 percent. Mean variance of housing prices ($VAR(P)$) is 4.9 percent, around a mean five year real growth $GROW(P)$ of 4.7 percent. The mean cell-level covariance ($COV(P, y_{ES202})$) is 0.6 percent, associated with a mean correlation CORR(P,y) of 0.29.

CORR(S& P,y) is the correlation between stock market returns (from CRSP's value weighted index) and cell income. Notably, the stock correlation measure is on average positive and significantly different from zero (although standard errors are biased by serial correlation, a problem difficult to solve with a small number of observations).

In stark contrast to the existing literature on housing and risk, I find similarly significant and typically positive correlations for stocks and housing prices $CORR(S\&P, P)$, with a mean of 0.21. The conventional view (as in Flavin and Yamashita [10]) that stock returns and house price increases are uncorrelated may again be premised on noisy short horizon estimation. In entering the stock market, workers must thus consider not only background income and price risk and stock market risk, but also considerable covariance between existing sources of wealth and stock market returns.

The income-price correlation results for particular industries (SICs) and MSA-SIC cells largely follow intuition. Some of these values are presented in Table 2. The largest income

Table 2: Income-House Price Correlations For Some Industries and Regions

Industry	MSA	$CORR(P, y_{ES202})$
Amusement and Recreation	Orlando	.64
Real Estate	National Average	.61
Auto Repair and Parking	National Average	.56
Building construction and General Contractors	National Average	.49
Security and Commodity Brokers, Dealers, etc.	New York	.44
Automotive Dealers	National Average	.42
Engineering, Accounting, Management, etc.	National Average	.41
All	National Average	.29
Petroleum Refining	Houston	.22
Transportation Equipment	Detroit	.18
Transportation Equipment	National	.12

Note: National average based on one observation per MSA-SIC cell is different from the individual-based averages in Table 4.

price correlation at the national level (taking averages over MSA-SIC cells' covariances with regional prices) belongs to the real estate industry, with a mean correlation of 0.61. Other large correlation industries are auto repair services and parking; automotive dealers; engineering, accounting research, management and related services; and building construction general contractors. Nationally, no industries have negative mean covariances.

The MSA-SIC cell that partly inspired this study, stock brokers in New York City, have the relatively large correlation of 0.44. Amusement and recreational workers in Orlando also have a predictably large correlation at 0.64. Auto workers (under the larger heading of transportation equipment industry workers) in Detroit have a correlation of 0.18, which is small relative to the overall national mean and relative to the Detroit MSA mean of 0.49 but large relative to the national transportation equipment industry mean of 0.12. Similar statistics apply to the petroleum industry in Houston.

4 Results

4.1 Effect of Covariance on Housing Purchases Combining the Extensive and Intensive Margins

The object of primary interest is the effect of income-price covariance, COV(P,y), on the value of housing owned, VALUE. Additional right hand side control variables labeled Z in equation (15) are the demographic and variance-covariance variables discussed above and summarized in Table 4.

Table 5 presents estimates of equation (15). Column (1) presents coefficient and standard error estimates of a regression of value on demographic variables, with MSA-SIC cell fixed effects and income interactions with MSA and SIC separately included but not reported. Recall that these controls absorb all effects of price levels and the distribution of prices both as a shared effect among all residents within a region and in interaction with income. Column (2) adds income-asset variance and covariance terms. Column (3) presents the first stage of the two stage least squares estimates, and column (4) the second stage IV estimates. The effects of some demographics are difficult to interpret because there are separate effects estimated for levels and for interactions with income. Education has a positive effect on demand both in levels and in interaction with income, as does family size. Being young (under 30, or between 30 and 40), black or Hispanic has a negative effect on housing owned both in levels and interacted with income.

When variance and covariance terms are added in column (2) of Table 5, we find the expected negative sign on income-price covariance. Variance of income and covariance with nominal interest rates and stock market returns also exert significant negative effects.

Turning attention to first stage IV results in column (3), we find that all ten instrumental variables are significant in the first stage COV(P,y) regression. If all or most of the demographic variables that are associated with homeownership were significantly correlated with $COV(P, y)$, we might suspect that the coefficient on $COV(P, y)$ picks up largely other effects. Despite large sample size, we find in column (3) that not all of the demographic covariates are significantly correlated with covariance.

Consistent with the estimated reliability ratio of $\frac{1}{4}$, the use of instrumental variables increases the estimated effect of covariance on housing investment by a factor of close to four. The coefficient on COV(P,y) increases in magnitude from -2.8 to -11.8 between OLS (specification (2)) and IV (specification (4)). Both coefficients are significant at a one percent confidence level. To interpret this coefficient, holding income constant, a one standard deviation increase in log covariance $COV(P, y_{ES202})$ of .01 would generate a decrease in housing purchases of just over ten percent of a household head's annual income. Alternatively, a one standard deviation increase in the level of covariance (including the income interaction) is 636, as reported in Table 4. Multiplying by the estimated coefficient of -11.8 implies a reduction in housing investment of approximately $7,500. A reasonable inference would be that housing purchases fall by approximately one bathroom per standard deviation increase in covariance.

4.2 Effect of Covariance on the Probability of Housing Purchases

To consider the effect of covariance on the choice between owning and renting housing, I evaluate average decisions and characteristics within MSA-SIC cells. While the effect of log income-price covariance on the intensive margin should surely grow with income, there is no obvious reason to think this would be the case regarding the own-rent decision. Since the correlation and log covariance measures are shared within cells, it is worthwhile in this setting to determine whether covariance effects can be detected at the cell level, at the sacrifice of over one million degrees of freedom. I treat each cell as an independent observation, which is justifiable only in the presence of MSA and SIC fixed effects. Standard errors are robust as throughout.

The cell level tenure choice analysis is presented in Table 6. The dependent variable is the fraction of household heads working in each of 7,396 MSA-SIC cells who rent their housing.

The effects of demographic variables are as expected, with cell mean age and variables associated with socioeconomic status generating positive effects on the probability of ownership. The effect of covariance is positive and significant in OLS, and larger but insignificant in IV estimation. The magnitude of the effect of covariance is quite small. Multiplying the IV coefficient of approximately 0.85 in the presence of demographic covariates by the standard deviation of log covariance (.01), yields a decrease of $\frac{85}{100}$ of one percent in the average fraction of homeownership within an MSA-SIC cell. Given the considerable variation that exists across cells in mean homeownership, this is a small effect. Further, the IV estimate is indistinguishable from zero (although this is predictable given the relatively small sample size, and the large number of fixed effects – 56 SIC codes plus 140 MSAs). The significance of the OLS effect is noteworthy, although measurement error tends to bias standard errors, as well as coefficients, downwards. Table 6 is thus weak evidence of a small effect of covariance on the "extensive" margin of tenure choice.

4.3 Effect of Covariance on the Value of Housing Owned, Conditional on Homeownership

Table 7 reports the estimated effect of the covariance of income on the "intensive" margin of housing value among homeowners only. The sample size is smaller than in the estimates reported in Table 5 because renters are excluded. The results are similar to the unconditional results reported in Table 5; youth and minority status are associated with small housing values. Income, education, whiteness and family size are associated with large housing assets. Covariance between income and prices has a significant negative effect on purchases and this effect is stronger when instrumental variables are used to overcome attenuation bias due to measurement error. The specification order is the same as in Tables 5 and 6. The IV coefficient in column (4) on COV(P,y) of -7.4 implies, holding income constant, that a one standard deviation increase in log income-price covariance $\mathrm{COV}(P, y_{ES202})$ would be associated with a reduction of approximately 7.4% of a year's income in housing value.

Alternatively, a one standard deviation increase in the covariance level would be associated with a reduction of approximately $4,700 in housing value conditional on ownership.

Following the model of section 2 and the results of Table 6, homeowners with large covariance values should be those with large idiosyncratic investment demand for housing. This would lead to a bias towards zero in our estimated effect of covariance. However, given the significant but small effect found on the own-rent decision, we anticipate only a small degree of bias due to selection. As discussed below, Heckman sample selection tests suggest that this is not an issue.

4.4 Nonlinearities, Sample Selection, and Liquidity Constraints

Table 8 summarizes results from a series of two step Heckman sample selection procedures that allow for nonlinear effects of $COV(P, y)$ on housing purchases and for sample selection. These estimates come from restriction of the IPUMS sample to small income ranges and small ranges of $COV(P, y_{ES202})$. These estimates are from regressions of the form:

$$Pr(OWN) = \Phi(\delta_0 COV(P, y_{ES202}) + \delta_1 Z^0);\ y_{ipums} \in [x, y],\ COV(P, y_{ES202}) \in [\gamma, \eta], \quad (26)$$

$$VALUE|OWN = \beta_0 + \beta_1 COV(P, y_{ES202}) + \beta_2 Z^1 + \epsilon;\ y_{ipums} \in [x, y],\ COV(P, y_{ES202}) \in [\gamma, \eta] \quad (27)$$

Equation (26) estimates the probability that an individual owns housing and equation (27) estimates the effect of covariates on the quantity of housing owned conditional on being a homeowner. Z^0 and Z^1 are different sets of covariates. Z^0 is the set of demographic controls used in the other estimates, but excludes income interactions and includes an additional variable, the length of time an individual has been in their current residence. Z^1 excludes length of time in the residence (to avoid collinearity) and includes the Inverse Mills ratio estimated in the probit (26). The results do not change significantly if the Inverse Mills ratio is excluded, suggesting that sample selection does not affect the results described above. This is consistent with the absence of strong effects of $COV(P, y_{ES202})$ on the extensive own-rent margin.

In each of equations (26) and (27), x and y are bounds on the income levels and γ and η are bounds on $COV(P, y_{ES202})$. I estimate 63 separate selection and conditional value equations. Separate estimates are generated for restrictions of the sample to income bins of width \$2,000 (income is measured continuously in IPUMS) from \$20,000 to \$60,000, encompassing approximately the 30th through 90th percentiles. These 21 estimates are undertaken three times each for the total of 63. The first time with the sample restricted to MSA-SIC cells with income-price covariances between the 25th and 35th percentiles (approximately 0 to .0013), the second time with the sample restricted to cells with covariances between the 45th and 55th percentiles (.0028 to .0045 covariance values), and the third time for cells with covariance between the 65th and 75th percentiles (.0062 to .0089). The percentiles are meant to capture a range of values while avoiding extreme values likely to reflect high measurement error.

The results in Table 8 are essentially consistent with the results presented above. The second column of Table 3 lists the percentile range of $COV(P, y_{ES202})$ values over which estimates are taken. Column (3) reports the mean coefficient on $COV(P, y_{ES202})$ on the dependent variable (either OWN or VALUE) across the 21 income bins between \$20,000 and \$60,000. There is little evidence of an effect of $COV(P, y_{ES202})$ on the own-rent decision, but the mean effect of covariance on conditional housing purchases is negative for all three value ranges of $COV(P, y_{ES202})$. Column (4) reports regression coefficients when the coefficient on $COV(P, y_{ES202})$ itself is regressed on the lower bound of each of the 21 income bins for which estimates are made. There is very limited evidence that the magnitude of the negative effect of $COV(P, y_{ES202})$ on housing purchases decreases with income. Almost none of the 63 individual coefficients estimated are significantly different from zero, and there is not enough information to determine curvature of the effect of covariance on housing purchases in either income or covariance.

The fact that the effect of covariance on value becomes more negative as income rises may be viewed as evidence that the results presented above are not driven by liquidity constraints, to the extent that liquidity constraints are typically associated with lower income households.

We would expect liquidity constraints to be associated with low income because demand for housing is inelastic in wealth (see, for example, Carliner [6]).

Further evidence that liquidity constraints, somehow correlated with income-price covariance, do not drive the result of decreasing housing purchases with increasing covariance comes from a restriction of the single sample regressions presented above in Tables 5 and 7. By restricting the sample to workers over age 45, we can largely rule out lifecycle-induced constraints. Under identical specifications, we find an effect of $COV(P, y)$ on $VALUE$ (counting rental as a zero purchase as in Table 5) of -2.05 (standard error .64) in OLS and -7.85 (2.55) in IV. These results are statistically significant, but slightly smaller than in the age-unconstrained sample reported above. In the conditional regressions, we again obtain similar results with the older sample and the unconstrained sample. Here the coefficients on $COV(P, y)$ are -1.98 (.48) and -6.88 (2.17), again very close to the age-unconstrained sample. We cannot rule out that the small difference between estimates relates to liquidity constraints, but an alternative explanation would be that cell-level income-price covariance is more important for younger workers with longer remaining careers.

5 Conclusions

Because housing is the most important asset, and labor income the most important source of wealth for most households, we expect intuitively that housing decisions will incorporate the desire to hedge against income risk. Putting some theoretical structure on the question of housing choice with risky prices and income, I find that households optimally purchase less housing on both the intensive and extensive margins as the covariance between housing prices and labor income increases. This theoretical prediction is borne out empirically. On the extensive margin, covariance has a negative effect on the probability of ownership, but significance is marginal and the magnitude of the effect quite small. On the intensive margin, the effect is clearer. On average, an increase of one standard deviation in covariance reduces housing investment by approximately $7,500, or, roughly speaking, one bathroom, including effects on both the extensive and intensive margins. An implication is that uninsurable labor income and housing prices, combined with non-diversification of housing investment, act to distort consumption and investment decisions substantially.

The results are interesting both because they extend our understanding of household financial risk and because they suggest that households, on average, are aware of these risks and take some measures to reduce risk. The data are consistent both with a large fraction of households making small housing investment modifications in response to joint income-price risk and with a small fraction of households making large modifications.

The theoretical and empirical results are interesting with respect to stock market behavior. Because homeowners are wealthier on average than renters, financial assets are concentrated in the hands of homeowners. On average, the incomes of these homeowners covary positively with housing prices. For homeowners considering the purchase of stock, there is thus background risk from income, from housing returns, and from the typically positive covariance of the two risks. Over long horizons, I find a positive correlation not only between stock market returns and labor income, but also between stock market returns and housing prices. The consequences for risk aversion over stock returns, and the welfare consequences of incremental investment in equities are worthy of further consideration.

None of the analysis presented would make sense if it were common for households to separate housing investment and consumption decisions. Robert Shiller and Allan Weiss have proposed that derivatives markets in regional housing prices might offset risk attributable to variability in capital gains on housing investment.[9] For similar reasons, Caplin et al. [5] propose financial instruments to allow homeowners to share home equity with broader markets. While the general equilibrium welfare effects of the introduction of such markets are ambiguous, the analysis suggests that such securities, if fairly priced, would have direct benefits for many households. Indeed, given the large average correlation found between income and prices, it appears that households might wish to hold short positions in regional price indices to smooth labor income across states of nature, independent of desire to smooth

[9] As in Shiller [19].

capital gains.

As a practical matter, most households directly hold few or zero non-housing assets, so that complete insurance against housing risk seems highly unlikely for most of the population. Given this and in light of the analysis presented here, proposals to remove the exemption of imputed rental services and the virtual exemption of capital gains on housing from taxation warrant further consideration. Berkovec and Fullerton [2] emphasize the attendant implicit risk sharing in housing prices. Assuming strictly positive nominal price changes (and no offsetting reduction in income taxes) a tax on housing capital gains would reduce the covariance between income and prices for homeowners and should hence proportionately reduce the significant consumption and investment distortion estimated above. Again, the general equilibrium welfare consequences are uncertain, but we might expect the presence of income-price covariance to augment the positive effects found by Berkovec and Fullerton. Heterogeneity in income-price covariances across households can be expected to complicate any such analysis.

References

[1] A. M. Aizcorbe, A. B. Kennickell, and K. B. Moore, Recent changes in u.s. family finances: Evidence from the 1998 and 2001 survey of consumer finances, *Federal Reserve Bulletin*, pages 1–32, January 2003.

[2] J. Berkovec and D. Fullerton, A general equilibrium model of housing, taxes and portfolio choice, *Journal of Political Economy*, 100(2):390–4429, 1992.

[3] J. Brueckner, Consumption and investment motives and portfolio choices of homeowners, *Journal of Real Estate Finance and Economics*, 15(2):159–180, 1997.

[4] J. Campbell and J. Cocco. Household risk management and mortgage choice. June 2001.

[5] A. Caplin, S. Chan, C. Freeman, and J. Tracy. *Housing Partnerships: A new approach to markets at a crossroads*. MIT Press, Cambridge, 1997.

[6] G. Carliner, Income elasticity of housing demand, *Review of Economics and Statistics*, 55(4):528–532, November 1973.

[7] K. E. Case and R. J. Shiller, The efficiency of the market for single-family homes, *The American Economic Review*, 79(1):125–137, March 1989.

[8] R. Chetty and A. Szeidl. Consumption commitments: Neoclassical foundations for habit formation. Manuscript, UC Berkeley, 2004.

[9] J. Cocco. Hedging house price risk with incomplete markets. September 2000.

[10] M. Flavin and T. Yamashita, Owner-occupied housing and the composition of the household portfolio over the life-cycle, *American Economic Review*, 2001.

[11] M. Fratantoni, Homeownership and investment in risky assets, *Journal of Urban Economics*, 44(1):27–42, 1998.

[12] Z. Griliches and J. Hausman, Errors in variables in panel data, *Journal of Econometrics*, 31(1):93–118, 1986.

[13] J. Heaton and D. Lucas, Portfolio choice and asset prices: The importance of entrepreneurial risk, *Journal of Finance*, 55(3):1163–1198, June 2000.

[14] J. V. Henderson and Y. Ioannides, A model of housing tenure choice, *The American Economic Review*, 73(1):98–113, March 1983.

[15] H. Lustig and S. V. Nieuwerburgh. Housing collateral, consumption insurance and risk premia: An empirical perpective. working paper 9959, NBER, 1993.

[16] A. Mas-Collel, M. Whinston, and J. Green. *Microeconomic Theory*. Oxford University Press, New York, 1995.

[17] F. Ortalo-Magné and S. Rady. Homeownership, low household mobility, volatile housing prices, high income dispersion. Manuscript, London School of Economics, University of Wisconsin, University of Munich, 2002.

[18] M. Piazzesi, M. Schneider, and S. Tuzel. Housing, consumption and asset pricing. UCLA, 2003.

[19] R. Shiller. *Macro Markets*. Clarendon Press, Oxford, 1993.

[20] T. Sinai and N. Souleles, Owner occupied housing as insurance against rent risk, *Quarterly Journal of Economics*, Forthcoming.

[21] G. Solon, Biases in the estimation of intergenerational earnings correlations, *Review of Economics and Statistics*, 71:172–174, 1989.

[22] S. Venti and D. Wise. Aging and housing equity. NBER Working Paper 7882, 2000.

[23] A. Vissing-Jorgenson. Towards and explanation of household portfolio heterogeneity: Nonfinancial income and participation cost structure. Working Paper, University of Chicago Department of Economics, 2000.

[24] T. Yamashita, Owner-occupied housing and investment in stocks: an empirical test, *Journal of Urban Economics*, 3:220–237, 2003.

[25] R. Yao and H. H. Zhang. Optimal consumption and portfolio choice with risky housing and stochastic labor income. University of North Carolina, Chapel Hill, 2001.

Table 3: Household Investment Income by Housing Tenure

	OWN	RENT
All	193,448	115,046
Have Investent Income	97,512	28,267
Mean Investment Income	2,470	534
Median Investment Income	10	0
Mean Home Value / Monthly Rent	135,451	470
Median Home Value / Monthly Rent	112,500	437

Notes: Data comes from 1990 US Census microdata (1 % sample). Values are for household heads with identifiable MSA-SIC cells and positive labor income.

Table 4: Summary Statistics

Variable (and description)	Obs	Mean	Std. Dev.	Min	Max
INC (Census income)	1,082,693	31,465	29,086	1.00	197,927
INC^2	1,082,693	1.84 Billion	4.9 Billion	1.00	39.2 Billion
BLACK (Indicator)	1,082,693	0.10	0.30	0.00	1
INC×BLACK	1,082,693	2,100	8,261.13	0.00	197,927
AGE (in years)	1,082,693	42	13.11	17	90
INC×AGE	1,082,693	1,354,855	1,446,073	30	17,600,000
FEMALE (Indicator)	1,082,693	0.28	0.45	0	1
INC×FEMALE	1,082,693	37,118	32,157	1.00	395,854
EDUC (Increasing function of years)	1,082,693	11.28	2.83	1.00	17
INC×EDUC	1,082,693	382,632	420,740	2.00	3,364,759
RENTER (Indicator)	1,082,693	0.37	0.48	0	1
VALUE|OWN	753,424	135,927	99,291	5,000	400,000
VALUE (Value, =0 for renters)	1,082,693	85,489	102,530	0.00	400,000
$CORR(P, y_{ES202})$	1,082,693	0.29	0.43	-1.00	1.00
$COV(P, y_{ES202})$	1,082,693	0.01	0.01	-0.33	0.63
$COV(P, y)$ $(=INC \times COV(P, y_{ES202}))$	1,082,693	191	635.56	-22,179	38,812
$VAR(y_{ES202})$	1,082,693	0.01	0.38	0.00	209
$INC \times VAR(y_{ES202})$	1,082,693	366	10,574	0.00	9,387,698
$GROW(y_{ES202})$ (mean growth)	1,082,693	1.04	0.06	0.43	7.64
$CORR(S\&P, y_{ES202})$	1,082,693	0.46	0.36	-1.00	1.00
$CORR(S\&P, P)$	1,082,693	0.21	0.40	-1.00	1.00
INC×$COV(S\&P, y_{ES202})$	1,082,693	580	1,098	-108,404	72,483
$COV(r, y_{ES202})$	1,082,693	-0.08	0.13	-4.56	2.20
INC×$COV(r, y_{ES202})$	1,082,693	-2,274	5,951	-368,530	200,791

Notes to Table 4 The level of observation is household heads in the 1990 US Census IPUMS 5 percent sample. y_{ES202} is mean MSA-SIC cell wage and percentage changes are observed between 1976 and 1999. P is the OFHEO house price index and percentage changes are observed between 1976 and 1999. Variance-covariance terms are based on five year overlapping horizons. r is average 30 year mortgage interest rates and S&P is the S&P index.

Table 5: Value of Housing Owned Regressed on Demographic and Covariance Characteristics

	(1)	(2)	(3)	(4)
DEPENDENT VARIABLE	VALUE	VALUE	COV(P,y)	VALUE
COV(P,y)		-2.821		-11.833
		(0.414)**		(1.569)**
INC	1.299	1.413	-0.032	1.184
	(0.099)**	(0.140)**	(0.0003)**	(0.150)**
AGE	1,507.48	1,507.62	0.227	1,508.50
	(27.100)**	(27.091)**	(0.082)**	(27.188)**
FEMALE	10,492.10	10,495.26	2.134	10,530.05
	(383.779)**	(383.490)**	-1.133	(384.033)**
FAMSIZE	3,713.56	3,719.63	1.232	3,738.27
	(99.007)**	(98.983)**	(0.307)**	(99.344)**
EDUC	3,341.63	3,338.33	0.662	3,345.85
	(56.600)**	(56.635)**	(0.169)**	(56.945)**
UNDER30	-1,250.76	-1,263.34	5.153	-1,243.29
	-766.594	-766.133	(2.329)*	-768.619
OVER30UNDER40	-7,567.92	-7,587.11	1.304	-7,603.95
	(484.178)**	(484.063)**	-1.472	(485.695)**
BLACK	-16,352.95	-16,352.05	0.97	-16,359.68
	(430.441)**	(430.652)**	-1.446	(433.134)**
NOTHISPANIC	19,062.77	18,921.86	-19.706	18,648.52
	(518.469)**	(519.036)**	(1.551)**	(522.990)**
$INC^2/10^{10}$	47,200	47,100	2.43	47,200
	(681)**	(681)**	1.58	(685)**
INC×AGE/1,000	-6	-6	$-2.43/10^7$	-6
	(1)**	(1)**	$(1.58/10^7)$**	(1)**
INC×FEMALE/1,000	-172	-0.172	-.1	-174
	(15)**	(15)**	(0.03)**	(16)**
INC×FAMSIZE/1,000	27	27	-.03	26
	(3)**	(3)**	(.007)**	(3)**
INC×EDUC/1,000	58	58	-.03	58
	(2)**	(2)**	(.004)**	(2)**
INC×UNDER30/1,000	-592	-592	-.2	-593
	(24)**	(24)**	(0.05)**	(24)**
INC×OVER30UNDER40/1,000	-6	-6	-.07	-6
	(13)	(13)	(0.03)*	(13)
INC×BLACK/1,000	-381	-381	-.02	-381
	(19)**	(19)**	(.05)	(19)**
INC×NOTHISPANIC/1,000	-19	-15	1	-6
	(21)	(21)	(0.04)**	(21)
INC×Grow(y)/1,000		-108	32	141
		(102)	(0.5)**	(117)
COV(S& P,y)		-1.482	0.043	-1.069
		(0.321)**	(0.001)**	(0.338)**
COV(R,y)		-0.199	0.007	-0.132
		(0.053)**	(0.00005)**	(0.056)*
INC×$VAR(y_{ES202})$		-0.022	0.001	-0.009
		(0.007)**	(0.00005)**	-0.017
INC×CORRN1			0.002	
			(0.00002)**	
INC×CORRN2			0.001	
			(0.00002)**	
INC×CORRN3			0.001	
			(0.00002)**	
INC×CORRN4			0.001	
			(0.00002)**	
INC×CORRN5			0.001	
			(0.00003)**	
INC×CORRN6			0.001	
			(0.00002)**	
INC×CORRN7			0.002	
			(0.00003)**	
INC×CORRN8			0.001	
			(0.00002)**	
INC×CORRN9			0.002	
			(0.00002)**	
INC×CORRN10			0.001	
			(0.00002)**	
Constant	0	0	0	0
	-70.957	-70.951	-0.246	-71.045
Observations	1,082,693	1,082,693	1,081,552	1,081,552
R-squared	0.38	0.38	0.64	N/A
Comment	OLS	OLS	IV Stage 1	IV Stage 2

Notes: Robust standard errors in parentheses, * denotes significant at 5%, ** at 1%. Indicator variables for MSA-SIC cells and income interactions with MSA and SIC fixed effects are included but not reported. Also included are level and income interactions of marital status indicators. VALUE is equal to the dollar value of household's housing unit if the household owner occupies, or zero if the household rents. INC is income. COV(P,y) is the covariance between income and house prices for a household head. UNDER 30 and OVER30UNDER40 refer to age. COV(S (R),y) is the covariance between income and stock market returns (nominal interest rates). CORRNX is the correlation between income and house prices in a household head's industry (SIC) in the Xth nearest MSA. These 10 variables are instrumental variables in specifications (3) and (4).

Table 6: Fraction of MSA-SIC Cell Workers Renting Housing Regressed on Income-Price Covariance and Other Demographic Characteristics

	(1)	(2)	(3)	(4)
DEPENDENT VARIABLE	RENTER	RENTER	COV(P,y_{ES202})	RENTER
COV(P,y_{ES202})		0.415		0.847
		(0.179)*		(0.957)
INC/10^{10}	-29,900	-29,600	-162	-0.000
	(4,020)**	(4,020)**	(292)	(0.000)**
AGE	-0.011	-0.011	0.000	-0.011
	(0.002)**	(0.002)**	(0.000)*	(0.003)**
FEMALE	-0.004	-0.001	-0.003	0.001
	(0.034)	(0.034)	(0.002)	(0.034)
FAMSIZE	-0.016	-0.016	0.001	-0.016
	(0.010)	(0.010)	(0.001)	(0.011)
EDUC	-0.010	-0.010	0.001	-0.010
	(0.004)*	(0.004)*	(0.000)*	(0.004)*
UNDER30	0.107	0.097	0.012	0.092
	(0.068)	(0.066)	(0.004)**	(0.069)
OVER30UNDER40	-0.031	-0.036	0.006	-0.039
	(0.042)	(0.042)	(0.003)*	(0.043)
BLACK	0.090	0.092	-0.005	0.094
	(0.034)**	(0.034)**	(0.003)	(0.034)**
NOTHISPANIC	-0.120	-0.120	0.002	-0.121
	(0.043)**	(0.043)**	(0.003)	(0.044)**
GROW(y)		-0.037	0.033	-0.051
		(0.037)	(0.003)**	(0.044)
$COV(S\&P, y_{ES202})$		-0.091	0.086	-0.127
		(0.103)	(0.007)**	(0.132)
COV(r,y_{ES202})		-0.025	-0.003	-0.023
		(0.017)	(0.001)*	(0.017)
VAR(y_{ES202})		0.001	-0.000	0.001
		(0.002)	(0.000)**	(0.001)
CORR5N1			0.002	
			(0.0004)**	
CORR5N2			0.001	
			(0.0004)**	
CORR5N3			0.000	
			(0.0004)	
CORR5N4			0.000	
			(0.0004)	
CORR5N5			-0.001	
			(0.0004)	
CORR5N6			0.000	
			(0.0004)	
CORR5N7			0.000	
			(0.0004)	
CORR5N8			0.001	
			(0.0004)*	
CORR5N9			0.001	
			(0.0004)*	
CORR5N10			0.0003	
			(0.0004)	
Constant	1.071	1.120	-0.060	1.145
	(0.160)**	(0.161)**	(0.010)**	(0.177)**
Observations	7396	7396	7396	7396
R-squared	0.62	0.63	0.29	0.62

Notes: Robust standard errors in parentheses. * Denotes significance at 5%, ** at 1%. All variables refer to MSA - SIC cell mean values. RENTER indicates that a household rents their housing. INC is income. COV(P,y_{ES202}) is the covariance between percentage changes in MSA-SIC cell mean income and house prices for a household head. UNDER 30 and OVER30UNDER40 refer to age. COV(S (R),y) is the covariance between income and stock market returns (nominal interest rates). CORRNX is the correlation between income and house prices in a household head's industry (SIC) in the Xth nearest MSA. These 10 variables are instrumental variables in specifications (3) and (4). Also included, but unreported are marital status, MSA and SIC dummies.

Table 7: Value of Housing Owned, Homeowners Only

	(1)	(2)	(3)	(4)
DEPENDENT VARIABLE	VALUE|OWN	VALUE|OWN	COV(P,y)	VALUE|OWN
COV(P,y)		-2.161		-7.389
		(0.338)**		(1.492)**
INC	0.949	1.094	-0.011	1.076
	(0.100)**	(0.134)**	(0.000)**	(0.135)**
AGE	566.593	566.876	0.182	567.452
	(27.730)**	(27.730)**	(0.112)	(27.791)**
SEX	14,504.872	14,502.568	0.154	14,517.946
	(486.006)**	(486.116)**	(1.837)	(487.030)**
FAMSIZE	1,175.403	1,179.257	1.101	1,194.964
	(110.969)**	(110.982)**	(0.447)*	(111.263)**
EDUC	4,679.676	4,678.734	1.037	4,687.120
	(68.043)**	(68.055)**	(0.254)**	(68.243)**
UNDER30	-279.718	-267.289	8.080	-260.138
	(856.006)	(856.173)	(3.493)*	(858.533)
OVER30UNDER40	-4,954.975	-4,960.520	1.996	-4,967.163
	(495.630)**	(495.624)**	(2.024)	(496.736)**
BLACK	-19,755.990	-19,761.873	2.903	-19,805.310
	(614.494)**	(614.445)**	(2.493)	(617.325)**
NOTHISPANIC	9,882.293	9,748.612	-26.561	9,532.603
	(688.768)**	(689.283)**	(2.611)**	(692.085)**
$INC^2/10^{10}$	-14,100	-14,100	-5.2	-14,200
	(608)**	(608)**	(1.98)**	(610)**
INC×AGE/1,000	-4	-4	-.005	-4
	(1)**	(1)**	(.002)*	(1)**
INC×FEMALE/1,000	-311	-311	-0.05	-312
	(15)**	(15)**	(0.04)	(15)**
INC×FAMSIZE/1,000	16	16	-0.02	15
	(3)**	(3)**	(0.009)*	(3)**
INC×EDUC/1,000	41	41	-0.03	41
	(2)**	(2)**	(0.006)**	(2)**
INC×UNDER30/1,000	-428	-428	-0.2	-428
	(24)**	(24)**	(0.08)**	(24)**
INC×OVER30UNDER40/1,000	-60	-60	-0.2	-60
	(11)**	(11)**	(0.03)*	(11)**
INC×BLACK/1,000	-249	-249	-0.1	-249
	(21)**	(21)**	(.1)	(21)**
INC×NOTHISPANIC/1,000	64	67	1	73
	(20)**	(20)**	(0.1)**	(20)**
INC×GROW(y)/1,000		-153	9	-131
		(92)	(0.3)**	(94)
COV(S& P,y)		-0.588	0.039	-0.378
		(0.285)*	(0.001)**	(0.299)
COV(R,y)		-0.126	0.008	-0.087
		(0.047)**	(0.000)**	(0.050)
INC×$VAR(y_{ES202})$		0.111	0.059	0.439
		(0.087)	(0.000)**	(0.136)**
INC×CORRN1			0.002	
			(0.00003)**	
INC×CORRN2			0.001	
			(0.00003)**	
INC×CORRN3			0.001	
			(0.00003)**	
INC×CORRN4			0.001	
			(0.00003)**	
INC×CORRN5			0.001	
			(0.00003)**	
INC×CORRN6			0.001	
			(0.00003)**	
INC×CORRN7			0.002	
			(0.00003)**	
INC×CORRN8			0.001	
			(0.00003)**	
INC×CORRN9			0.002	
			(0.00003)**	
INC×CORRN10			0.001	
			(0.00003)**	
Constant	-12,526.508	-12,527.645	-0.247	-12,545.816
	(81.893)**	(81.890)**	(0.396)	(82.035)**
Observations	661,934	661,934	660,987	660,987
R-squared	0.31	0.31	0.66	N/A

Notes: Robust standard errors in parentheses, * denotes significant at 5%, ** at 1%. Indicator variables for MSA-SIC cells and income interactions with MSA and SIC fixed effects are included but not reported. Also included are level and income interactions of marital status indicators. VALUE is equal to the dollar value of household's housing unit if the household owner occupies, or zero if the household rents. INC is income. COV(P,y) is the covariance between income and house prices for a household head. UNDER 30 and OVER30UNDER40 refer to age. COV(S (R),y) is the covariance between income and stock market returns (nominal interest rates). CORRNX is the correlation between income and house prices in a household head's industry (SIC) in the Xth nearest MSA. These 10 variables are instrumental variables in specifications (3) and (4).

Table 8: Summary of Heckman Sample Selection Estimates For Different Income and Covariance Bins

Dependent Variable	Covariance Percentiles	Mean Coefficient on $COV(P, y_{ES202})$	Regression Coefficient of Income on Coefficient
(1)	(2)	(3)	(4)
OWN	25-35	28.08	-.0002
OWN	45-55	345	.042
OWN	65-75	-159	-.018
VALUE	25-35	-1,304,500	-192
VALUE	45-55	-167,994	82
VALUE	65-75	-398,376	-55

Notes: This table summarizes Heckman sample selection estimates as described in equations (26) and (27). Coefficient means in the third column are taken over 21 separate confinements of the data to income ranges of width \$2,000 from 20,000 to 60,000. The fourth column reports an OLS estimate of the effect of the lower bound of the income bin on the estimated coefficient reported in the third column, with 21 observations per row.