Supplemental Material (not copyedited or formatted) for: Dennis Epple, Luis Quintero, Holger Sieg. 2020. "A New Approach to Estimating Equilibrium Models for Metropolitan Housing Markets." Journal of Political Economy 128(3). DOI: https://doi.org/10.1086/704755.

A New Approach to Estimating Equilibrium Models for Metropolitan Housing Markets: Online Appendices

Dennis Epple

Carnegie Mellon University and NBER

Luis Quintero

Johns Hopkins University

Holger Sieg

University of Pennsylvania and NBER

May 23, 2019

A A Constructive Proof of Identification Based on an Example With Closed Form Solutions

Here we consider a simplified version of our model with continuous quality measures and no unobserved heterogeneity among households. To obtain a closed form solution for the equilibrium pricing function, we impose additional functional form assumptions.

Assumption 1 *Income and housing are distributed generalized log-normal with location parameter (GLN4).*[1]

$$\begin{aligned} \ln(y_t) &\sim GLN4(\mu_t, \sigma_t^{r_t}, \beta_t) \\ \ln(v_t) &\sim GLN4(\omega_t, \tau_t^{m_t}, \theta_t) \end{aligned} \tag{1}$$

We will show below that these functions are sufficiently flexible to fit the housing value and income distributions in the metro areas and time periods that we consider in the empirical analysis. Imposing the restriction that $r_t = m_t$ permits us to obtain a closed-form mapping from house value to income. We then establish that the further assumption that $\theta_t - \beta_t$ is time invariant permits us to obtain a closed-form solution to the hedonic price function.

Proposition 1 *If $r_t = m_t \ \forall t$, the income housing value locus is given by the following expression:*

$$y_t = A_t \, (v_t + \theta_t)^{b_t} - \beta_t \tag{2}$$

with $a_t = \mu_t - \frac{\sigma_t}{\tau_t}\omega_t$, $A_t = e^{a_t}$, and $b_t = \frac{\sigma_t}{\tau_t}$.

For our discussion of identification below, it is useful to note that all of parameters of the sorting locus, $a_t = \mu_t - \frac{\sigma_t}{\tau_t}\omega_t$, $A_t = e^{a_t}$, $b_t = \frac{\sigma_t}{\tau_t}$, and θ_t can be estimated directly

[1] The four-parameter distribution for income simplifies to the standard two-parameter lognormal when the location parameter β_t equals zero and the parameter r_t equals 2.

from the data. In addition, it will be useful below to note that if $b_t > 1$, this function is convex. To obtain a closed form solution for the equilibrium price function, we adopt the following functional form for household preferences.

Assumption 2 *Let utility given by:*

$$U = u_t(h) + \frac{1}{\alpha}\ln(y_t - v_t(h) - \kappa) \tag{3}$$

with $u_t(h) = \ln(1 - \phi(h+\eta)^{\gamma})$*, where* $\alpha > 0$*,* $\gamma < 0$*,* $\phi > 0$*, and* $\eta > 0$.[2]

In addition to yielding a closed-form solution for the hedonic price function, this utility function proves to be relatively flexible in allowing variation in price and income elasticities. Given this parametric specification of the utility function, we have the following result:

Proposition 2 *If* $b_t > 1$ *and* $\kappa = \theta_t - \beta_t \;\; \forall t$*, the hedonic equilibrium pricing function is unique and given by:*

$$v_t = \left(A_t\left[1 - \frac{(1-\phi(h+\eta)^{\gamma})^{\alpha(b_t-1)}}{e^{c_t}}\right]\right)^{\frac{1}{1-b_t}} - \theta_t \tag{4}$$

for all $h > (\frac{1}{\phi})^{\frac{1}{\gamma}} - \eta$*.* c_t *is the constant of integration.*

Note that parameters A_t, b_t, θ_t can be estimated directly from data for income and house rent distributions. We show these are sufficient for identification of the utility function parameters. First consider the normalization $v_t(h) = h$. Assume that the constant of integration c_t is set to zero in the baseline period. Hence, the equilibrium hedonic pricing function is given by:

$$v_t = \left(A_t\left[1 - [1 - \phi(h+\eta)^{\gamma}]^{\alpha(b_t-1)}\right]\right)^{\frac{1}{1-b_t}} - \theta_t \tag{5}$$

[2]This utility function requires the following two conditions be satisfied $1 - \phi(h+\eta)^{\gamma} > 0$ and $y_t - v_t - \kappa > 0$.

Setting

$$\alpha = \frac{1}{b_t - 1} \tag{6}$$

implies

$$v_t = (A_t\left[1 - [1 - \phi(h+\eta)^\gamma]\right])^{\frac{1}{1-b_t}} - \theta_t \;=\; (A_t \phi(h+\eta)^\gamma)^{\frac{1}{1-b_t}} - \theta_t \tag{7}$$

Setting

$$\phi = \frac{1}{A_t} \tag{8}$$

implies

$$v_t = ((h+\eta)^\gamma)^{\frac{1}{1-b_t}} - \theta_t \tag{9}$$

Setting

$$\gamma = 1 - b_t \tag{10}$$

implies

$$v_t = (h+\eta) - \theta_t \tag{11}$$

Finally, setting

$$\eta = \theta_t \tag{12}$$

implies.

$$v_t = h \tag{13}$$

That establishes identification of the parameters of the utility function. Hence the normalizations that $c_t = 0$ and that $v_t(h) = h$ are sufficient to identify the parameters of the utility function.

The price equation in any other period $t+s$ is then given by:

$$v_{t+s} \;=\; \left(A_{t+s}\left[1 - \frac{(1-\phi(h+\eta)^\gamma)^{\alpha(b_{t+s}-1)}}{e^{c_{t+s}}}\right]\right)^{\frac{1}{1-b_{t+s}}} - \theta_{t+s} \tag{14}$$

The assumption of constant utility across time then implies that $v_{t+s}(h)$ is identified by the parameters b_{t+s}, A_{t+s}, and θ_{t+s} and the normalization that $v_{t+s}(0) = 0$. Note that we need the last normalization to determine the constant of integration c_{t+s}.

B Accounting for Owners' Equity

To account for the role of housing wealth in the earnings of households, we calculate the imputed owner-occupied rental equivalent that homeowners pay to themselves and add it to their income. The corresponding rental income of the homeowners is proportional to the equity they have on the unit they occupy. We calculate the equity for each household by subtracting the outstanding balance of the mortgage(s) from the house value. Then, we transform it to the implied rental income using the original user cost, and finally add this to the current income.

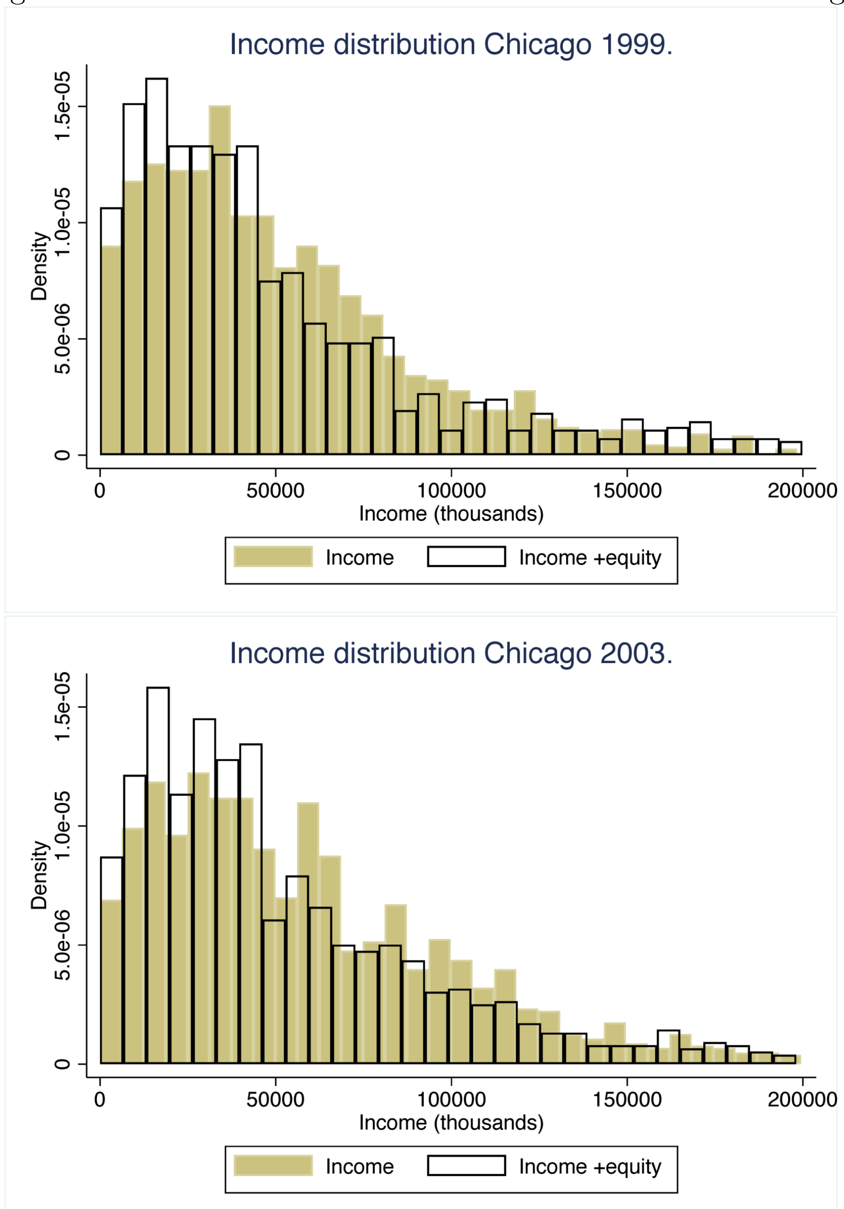
The AHS does not ask respondents directly about their current equity. However, they produce code that uses their mortgage(s) information to calculate the outstanding principle.[3] The data codebook notices the issue of top-coding in both housing values and debt in calculating this amount. However, it also clarifies this is the only way of calculating equity with the AHS data. Note that the estimates we obtain of owners' mortgage debt equity are broadly consistent with the ones reported in Li and Goodman (2016), while the outstanding debt we calculate, as a proportion of house value, are similar to those reported in the Flow of Funds Accounts Statistical Release for the corresponding periods calculated by the Federal Reserve Board of Governors.

Using the new comprehensive income measure described above, we follow the iterative procedure described in section 4.2, re-estimating the rent-to-value function, imputing again the rental income for owners, and iterating until convergence. We find that we reach convergence fast. Table 1 shows summary statistics for the calculated outstanding debt as percentage of the homeowner's unit value and the equity.

The initial income distribution and the distribution that accounts for owners' housing equity are plotted in figures 1 below for 1999 and 2003 respectively. Note that

[3] Census produces a document called *AHS Table Specification* that includes SAS code used to produce publications from HUD USER. The section for the variable OTPIN includes the code to calculate the outstanding principle or debt balance.

Figure 1: Difference in Income Distributions Functions: Chicago



Supplemental Material (not copyedited or formatted) for: Dennis Epple, Luis Quintero, Holger Sieg. 2020. "A New Approach to Estimating Equilibrium Models for Metropolitan Housing Markets." Journal of Political Economy 128(3). DOI: https://doi.org/10.1086/704755.

Table 1: Equity

	Mean	Median
Mortgage debt[a] 1999	54%	55%
Mortgage debt 2003	54%	54%
Equity[b] 1999	94,464	79,524
Equity 2003	104,891	91,374

income increases for the majority of owners since most have positive equity in the house. However, there are also some decreases in income due to negative owners' equity. The distribution does not change dramatically.

C Preference Parameter Estimates for the Joint NYC - CHI Model

The estimates of the parameters of the utility functions of the different household types for the joint NYC and Chicago model are summarized in Table 2. Overall we find that the results are similar to the one we obtained when we just used the Chicago subsample. We hence conclude that preferences for housing are similar in both metropolitan areas. The main difference is that we obtain slightly smaller standard errors for the structural parameters of the utility function which suggests that there are some efficiency gains from pooling across metropolitan areas.

Tables 3 report the estimates for the unobserved types and the implied probabilities for the observed types. We find that a model with five unobserved types fits the data well.

Table 2: Preference Parameter Estimates: CHI-NYC

	α	ϕ	η	γ
i_1	1.33	9.31	0.45	-1.91
	(0.42)	(2.00)	(0.08)	(0.14)
i_2	2.34	3.23	1.56	-1.12
	(0.35)	(1.51)	(0.17)	(0.08)
i_3	2.77	2.13	4.78	-1.13
	(0.45)	(0.81)	(0.15)	(0.13)
i_4	1.01	7.32	1.11	-1.43
	(0.50)	(0.51)	(0.15)	(0.11)
i_5	1.87	5.65	2.38	-1.23
	(0.60)	(2.01)	(0.21)	(0.10)

Table 3: Implied probability of types: CHI-NYC

π_{ikt}	x_1	x_2	x_3	agr
i_1	0.096	0.024	0.013	0.134
i_2	0.011	0.299	0.027	0.337
i_3	0.014	0.017	0.147	0.178
i_4	0.100	0.057	0.030	0.187
i_5	0.035	0.078	0.050	0.164