Replication

Guilherme Cota

2023-10-26

1. Identification and Estimation

1.1. A Parameterization of the Model

To identify and estimate the model, it is necessary to introduce flexible parameterizations of the key functions of interest.

- **The Utility Function**

The utility provided by housing quality h for household type i at each period t is given by

$$U = u_i(h) + \frac{1}{\alpha_i} ln(y_t - v_t(h))$$

with

$$u(h) = ln(1 - \phi_i(h + \eta_i)^{\gamma_i})$$

where $\alpha_i > 0$, $\gamma_i < 0$, $\phi_i > 0$, and $\eta_i > 0$.

This utility function requires that the terms in inside the logarithm be positive and the authors impose these structural constraints in estimation. They also assume that the parameters of the utility function are time invariant.

This utility function gives rise to flexible forms for the price and income elasticity equations. It also allows us to obtain a closed-form solution for the income cutoffs.

- **The Law of Motion for Housing Supply**

The law of motion for the fraction of houses in each quality bin h_{jt} is given by

$$r_{jt} = A_t r_{jt-1} \left(\frac{V_{jt}}{V_{jt-1}} \right)^{\zeta} \forall j, t$$

where A_t is a scalar that guarantees that the fractions of housing types sum to one.

1.2. Rent-To-Value Functions

In the data, we observe rents for rental units and values for owner-occupied units, we never observe rents and values for the same unit. As a consequence, we have to impose some assumptions to identify the rent-to-value functions.

We assume that all households are indifferent between renting and owning.

The expected rent for a household of type (x, y) is given by

$$v_t(x, y) = \sum_i p_{it} v_t(h_{it}(y))$$

and the expected housing value is given by

$$V_t(x, y) = \sum_i p_{it} V_t(h_{it}(y))$$

The authors estimate the conditional expectations above using a nonparametric estimator. For each type x, fix quality at h and find the value of y such that $v_t(h) = v_t(x, y)$ and define $V_t(h) = V_t(x, y)$.

This matching algorithm defines a mapping, $v_t^x(V)$, for each observed household type x. Equilibrium requires that these mappings be the same for all x.

Hence, we define the rent-to-value function, $v_t(V)$, as the following projection:

$$\min_{v_t(V)} \sum_x s_t(x) \int ||v_t(V) - v_t^x(V)|| dV$$

For the discretized version of the model

$$\min_{v_{1t}, \ldots v_{jt}} \sum_x s_t(x) \sum_j (v_{jt} - v_{jt}^x)^2$$

Proposition: The rent-to-value function $v_t(V)$ and its inverse function $V_t(v)$ are nonparametrically identified.

Having identified the rent-to-value functions, we can convert values for owner-occupied houses into rents and hence construct the complete rental distribution for each market and each time period.

1.3. Latent Housing Quality and Prices

Housing quality is ordinal and latent, so there is no intrinsic unit of measurement for housing quality. We can use an arbitrary monotonic transformation of h and redefine the utility function and supply function accordingly. We can normalize housing quality by setting $h = v_1(h)$.

The normalization of housing quality is a useful way to generate initial conditions for our model that are consistent with observed outcomes.

1.4. A Method-of-Moments Estimator

The model use the following estimation algorithm:

(1) Estimate for each t the rent-to-value function, $v_t(V)$, and its inverse, using a nonparametric matching estimator.

(2) Use the rent-to-value function to impute rents for owner-occupied units, obtain the market distribution for rents, and compute the joint distributions for income and rent conditional on x, denoted by $F_t^N(y, v|x)$.

(3) Discretize the rent distribution in period 1 into J intervals indexed by j, and normalize housing quality to obtain h_j and r_{j1} for all j.

(4) Choose a vector of structural parameters, denoted by v, that includes the parameters of the type distribution. Compute the type-specific income distributions $F_{it}(y)$.

(5) Solve for the implied equilibrium prices in all periods t > 1.

(a) Guess values of v_{jt}.

(b) Calculate the implied income cutoffs $\hat{y}_{ijt}$ such that

$$U_i(h_j, \hat{y}_{ijt} - v_{tj}) = U_i(h_{j+1}, \hat{y}_{ijt} - v_{tj+1}), \forall t, it$$

(c) Calculate the implied demands:

$$H^d_{ijt}(v_t) = F_{it}(\hat{y}_{ijt}) - F_{it}(\hat{y}_{ij-1t}), \forall t, it$$

(d) Calculate the supplies:

$$r_{jt}(v_t) = A_t r_{jt-1} \left(\frac{V_{jt}(v_{jt})}{V_{jt-1}(v_{jt-1})} \right)^{\zeta}, \forall j, t > 1$$

(e) Check whether equilibrium holds:

$$\sum_x s_t(x) \sum_{i=1}^{I} p_{it}(x) H^d_{ijt}(v_t) = r_{jt}(v_t), \forall j, t > 1$$

(f) Repeat until the solution of equilibrium prices has been found for each time period.

(6) For each unobserved type i, compute the predicted joint $F^{\theta}_{it}(y, v)$ as well as $F^{\theta}_{it}(y, v|x)$.

(7) Form orthogonality conditions that are based on the difference between the observed joint distributions of income and rents, denoted by $F^N_t(y, v|x)$, and their predicted counterparts, $F^{\theta}_{it}(y, v|x)$. Specifically, we use quantiles of the marginal distributions and correlations between y and v conditional on x. Evaluate the objective function $Q^N(\theta)$.

(8) Update θ until $Q^N(\theta)$ is minimized.

The objective function $Q^N(\theta)$ minimizes moments that measure the difference between quality supply and demand, the difference between the implied and the observed aggregate rent distributions, and the difference between observed and estimate income and rent correlations.

1.5. Some Remarks

Before explaining the results, we need to make three observations regarding the replication.

The MATLAB code takes approximately 25 days to run, and based on our estimates, it would take 8 times longer in R. Therefore, we had to significantly reduce the number of iterations, from 15,000,000,000 to 60,000,000. As a result, the R code ran in approximately 55 hours.

The code is in R and not in R Markdown because it is too large for compilation. It is divided into 33 auxiliary functions used to estimate the model. There are two main functions that estimate the model for 1 (a2Estimate) and 2 (a3Estimate) cities. There is also a code (a0Master) that calls all 35 functions and performs all the estimations. It takes approximately 55 hours to run.

The author said consider two specifications that differ by the income measure. The first model is based on the commonly used income measure reported by the AHS. The second model uses a comprehensive income measure that also accounts for owners' equity, adding imputed rental income to the first measure. In the replication package, there is only an estimation of the first specification, and this is the one we are going to replicate.