Macro Labor with Micro Data: BPP, Kaplan-Violante and the Bewley Model

Zachary Orlando & Cutberto Frias Sarraf

Part 1: Data

PSID Sample Selection and Demographics

We imposed the following sample selection criteria on the PSID dataset:

1. We drop all indiduals from the SEO oversample.
2. We restrict the sample to only individuals who are considered the heads of their household.
3. To fit with the model, we further consider only observations of individuals who are aged between 25 and 59.
4. Observations must be between years 1978 and 1997 of the PSID.
5. After applying these restrictions, we then used a minimum income threshold of $3500 for an observation to be counted. Households must satisfy this threshold for at least five years (so that we can measure three changes in growth rates of income for each household in my sample) or be dropped entirely.
    - $3500 was approximately the first percentile of the household income distribution.

Table 1: Summary Statistics by Year

Year	Income (000s)	Age	Gender	Education	Working	White	Black	HH Size	Num Children	Observations
1978	39.96	37.09	0.13	12.88	0.92	0.89	0.08	3.29	1.30	1820
	(19.89)	(9.08)	(0.34)	(2.54)	(0.27)	(0.32)	(0.27)	(1.54)	(1.27)	
1982	35.61	38.39	0.15	12.92	0.88	0.89	0.08	3.10	1.15	2351
	(24.71)	(9.90)	(0.35)	(2.44)	(0.33)	(0.31)	(0.27)	(1.43)	(1.19)	
1986	40.38	38.68	0.15	13.40	0.89	0.89	0.08	3.05	1.12	2512
	(27.02)	(9.45)	(0.36)	(2.48)	(0.31)	(0.32)	(0.27)	(1.42)	(1.18)	
1990	43.99	39.27	0.16	13.55	0.90	0.88	0.08	3.02	1.10	2606
	(36.70)	(8.80)	(0.37)	(2.35)	(0.30)	(0.32)	(0.27)	(1.42)	(1.19)	
1994	46.00	41.35	0.14	13.60	0.91	0.88	0.08	3.06	1.10	2457
	(41.44)	(8.09)	(0.35)	(2.31)	(0.29)	(0.32)	(0.26)	(1.43)	(1.18)	
1997	51.27	43.55	0.13	13.67	0.91	0.89	0.07	3.06	1.08	2185
	(39.80)	(7.48)	(0.34)	(2.28)	(0.29)	(0.32)	(0.26)	(1.43)	(1.18)	

* Mean and (SD).
Notes: Post-Government income uses the NBER TAXSIM estimates and is in constant 1997 dollars. Gender = 0 for Man. Education is in years.

We report all demographics at the level of the household head. We used Post-Government TAXSIM household income as our measure of income, because it continues past 1991. Furthermore, as income is measured in current-year dollars, we adjust for inflation using CPI so that all income is in constant 1997 dollars.
Household heads in our PSID sample are primarily middle-age white males. The average age in our sample trends upwards, which is unsurprising considering households which enter the sample later may be unable to satisfy our income filter five times.

Constructing κ_t

We seek to estimate the predictable component of income, κ_t, and the residual component, y_{it}, from the below model:

$$
\begin{aligned}
\log Y_{it} &= \kappa_t + y_{it} \\
y_{it} &= P_{it} + \varepsilon_{it} \\
P_{it} &= \rho P_{i,t-1} + \zeta_{it} \\
\zeta_{it} &\sim N(0, \sigma_\zeta) \\
\epsilon_{it} &\sim N(0, \sigma_\epsilon) \\
\rho &= 0.97
\end{aligned}
$$

To construct the predictable component of income, we regressed log household income on dummies for age and year of birth. κ_t in the model is the mean fitted value by age. The residual vector from this regression is y_{it} which we use to estimate the variance of persistent and transitory income.

Figure 1:

Additional demographic variables that match, for instance, the procedure in BPP, such as employment status, state, years of education, white or black, household size and number of children led to no noticeable change in the estimated predictable component.

We can see that lifecycle post-government income is roughly hump-shaped and peaks a little after age 45.

Estimating the variance of the persistent and transitory income shocks

After obtaining κ_t and y_{it}, we estimated the parameters, σ_ζ and σ_ϵ, of the model using the Generalized Method of Moments (GMM).

We used the BPP moment conditions to identify the parameters:

$$\mathrm{Cov}(\Delta\tilde{y}_{it}, \Delta\tilde{y}_{i,t+1}) = -\rho * \sigma_\epsilon \tag{1}$$

$$\mathrm{Cov}(\Delta\tilde{y}_{it}, \rho^2\Delta\tilde{y}_{i,t-1} + \rho\Delta\tilde{y}_{i,t} + \Delta\tilde{y}_{i,t+1}) = \rho * \sigma_\zeta \tag{2}$$

Where $\tilde{y}_{it} = y_{it} - \rho y_{it}$.

Table 2 shows that our results are roughly similar to the BPP estimates.

	σ_ζ	σ_ϵ
Estimate	0.0228	0.03863
Std. Error	0.0049	0.0049

Table 2: Parameter Estimates

Other moments in theory should be useful for identifying the parameters if we believe our model. For example, for $\rho < 1$, the variance and covariance of log income should be useful:

$$\mathrm{Var}(y_{it}) = \frac{\sigma_\zeta}{1-\rho^2} + \sigma_\epsilon \tag{3}$$

$$\mathrm{Cov}(y_{it}, y_{i,t+1}) = \frac{\rho\sigma_\zeta}{1-\rho^2} \tag{4}$$

However, when we incorporated these additional moment conditions into our estimation using two-step GMM, the parameter estimate for the variance of permanent income was too small to be useful for this exercise (0.00083). The estimate for the variance of transitory income seemed reasonably close to the one employed in Kaplan-Violante (0.05873).

We infer this is because the additional moments are much more sensitive to the true value of ρ, particularly close to $\rho = 1$.

We further checked using just the covariance and variance of $\Delta(y_{it})$ (which again identify the parameters under the model). These conditions produced insurance estimates that were reasonably close to the BPP ones.

Part 2: Model

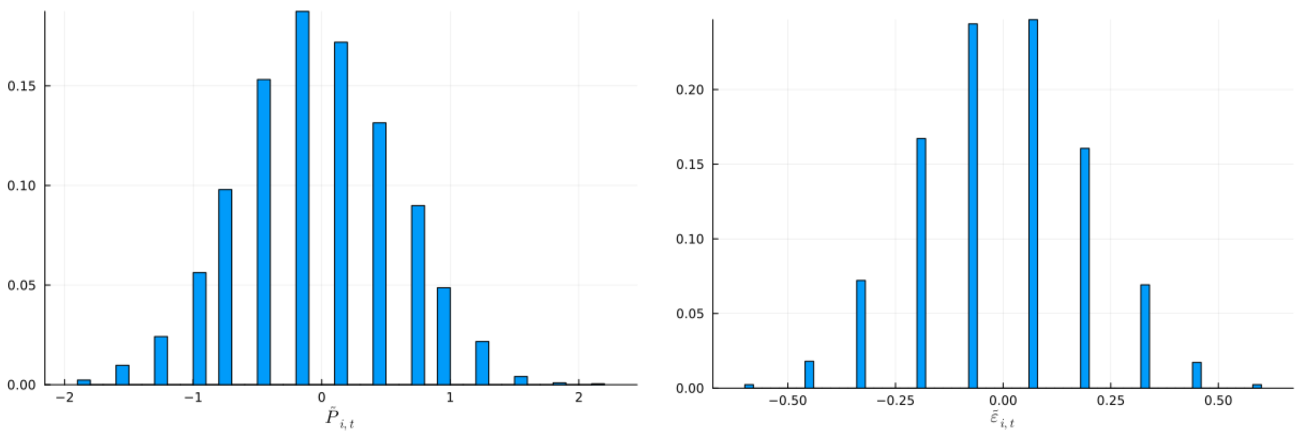
Discretizing transitory and persistent income:

We used the Rouwenhorst method to discretize the income processes. Noticing that transitory and persistent income are independent processes, we discretized separately.

We used 20 grid points for the persistent process and 10 for the transitory process.

The distribution of the initial persistent component, $\tilde{P}_{i0}$, is constructed by associating each grid point of $\tilde{P}_{it}$ with the change in the cumulative CDF of a $N(0, \sigma_{\zeta 0})$ random variable at the grid point's value from the value of the prior grid point (with the last value of the grid taking all remaining density). Just like in Kaplan-Violante, $\sigma_{\zeta 0}$ is assumed to be 0.15 .

Figure 2:



To check, we simulated the stationary distributions and they look reasonable.

Solving and Simulating a Bewley Model

Using the discretized permanent and transitory income processes along with the deterministic lifecycle income path estimated in part 1, we solved the decision problem of a Bewley model with C.R.R.A utility and relative risk aversion of $\gamma = 2$. We solved the model with a **zero-borrowing constraint** (ZBC). The recursive representation of our model is:

$$V(a, \epsilon_{it}, P_{it}) = \max_{c,a'} \{u(c) + \beta E [V(a', \epsilon_{i,t+1}, P_{i,t+1}) \mid \epsilon_{it}, P_{it}]\} \quad \text{t = 1, 2,..., 35}$$

subject to the budget constraint:

$$a' = (1 + r)a + Y(\epsilon_{it}, P_{i,t}) - c, \quad a' \geq 0$$

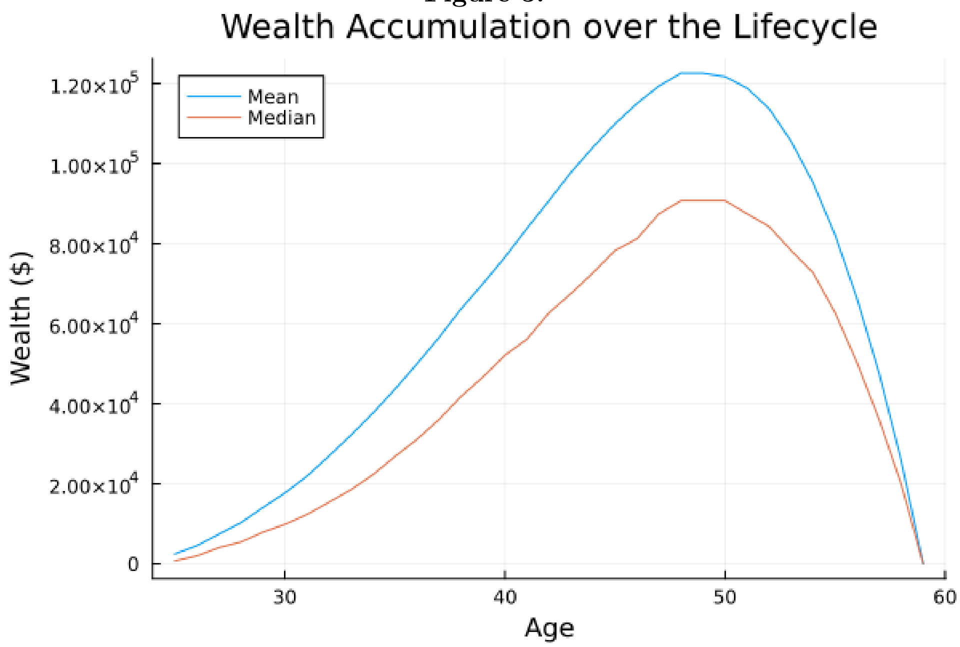
and the income process:

$$\log(Y_{it}) = \kappa_t + \epsilon_{it} + P_{it}$$
$$P_{it} = \rho P_{i,t-1} + \zeta_{it}$$
$$\zeta_{it} \sim N(0, \sigma_\zeta),\ \epsilon_{it} \sim N(0, \sigma_\epsilon),\ P_{i0} \sim N(0, \sigma_{\zeta 0}),$$

Where when solving we replace P_{it}, P_{i0} and ϵ_{it} with our discretized processes $\tilde{P}_{it}$, $\tilde{P}_{i0}$ and $\tilde{\epsilon}_{it}$.

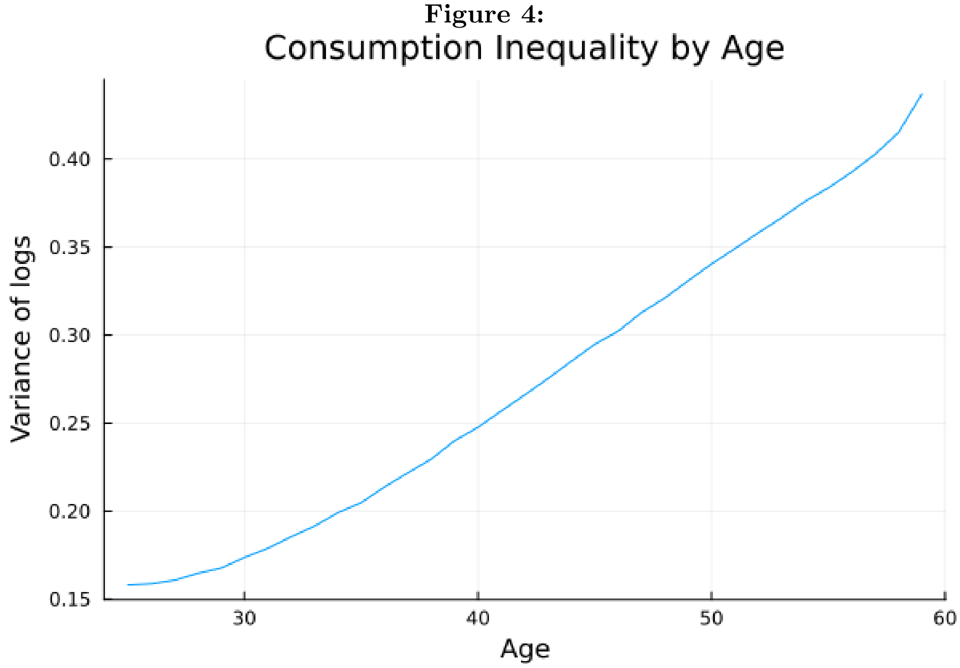
Similar to Kaplan-Violante we used 500 exponentially-spaced grid-points for assets with a maximum level of \$1,000,000. We show that this level is appropriate in Figure 4. below.

Figure 3:



In our economy, individuals steadily accumulate wealth, although the peak at just before age 50 appears to be more hump-shaped than in Kaplan-Violante. This could be due to the absence of retirement in our model. Furthermore, households dis-save much more rapidly.

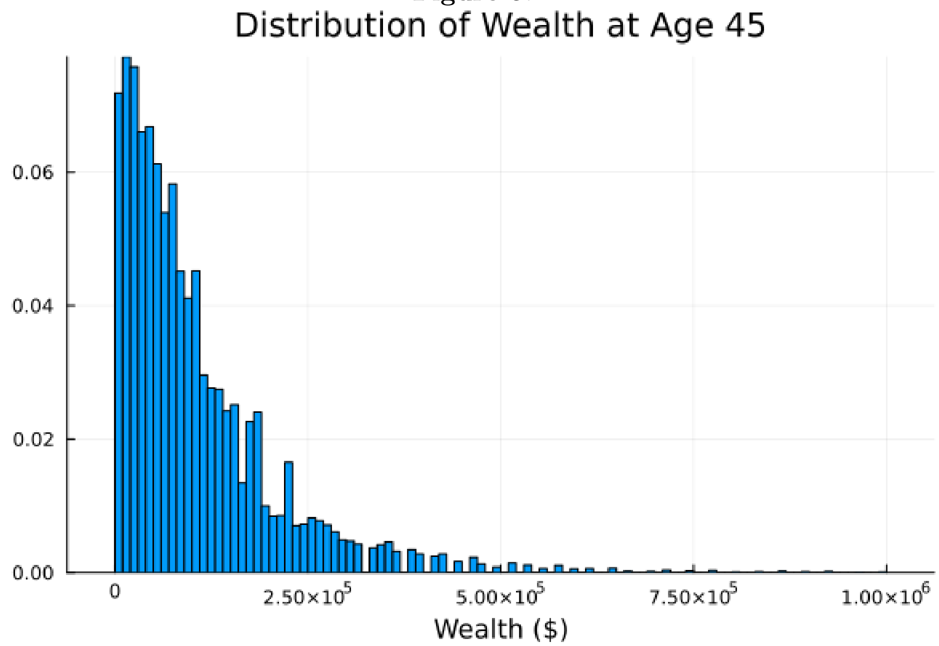
Figure 4:



Consumption inequality rises roughly log-linearly as persistent shocks cumulate, with initial variance in logs similar to the variance in the initial persistent component. This is very similar to the zero borrowing constraint case in Kaplan-Violante.

We also verified that our choice of maximum wealth was acceptable for an initial $a = 0$ problem by plotting the wealth histogram across 20,000 simulations at age 45 (around the age of highest mean wealth).

Figure 5:



Estimating the insurance coefficients

Our results in Table 3. are broadly similar to those in Kaplan-Violante.

- "Data BPP" estimates of a 5% pass-through of transitory shocks and a 64% pass through for permanent shocks overestimates the degree of insurance (by roughly 13% for both shocks).
- Comparing "Model BPP" to the "True Model" coefficients, we see that BPP is much less biased for measuring insurance against transitory shocks than persistent shocks. Furthermore, it tends to undershoot the degree of insurance for a persistent shock and if anything overshoot for a transitory shock.

Parameter	BPP Model	True Model
α^{ϵ}	0.8253	0.8162
α^{ζ}	0.2323	0.2681

Table 3: Parameter Estimates

The primary difference we noticed was that the coefficients are lower for the persistent shock than reported for $\rho = 0.97$ with a ZBC in Kaplan-Violante. Perhaps this is due to the lack of retirement in the estimated model. Individuals have less of an incentive to save and thus are less insured against shocks when they arrive.