ECON-810: Problem Set 1 Solutions

Fernando de Lima Lopes*, Stefano Lord Medrano† and Yeonggyu Yun‡

University of Wisconsin-Madison

February 1, 2022

Part 1: Data

Using the PSID data in the Dropbox folder, estimate the variance of temporary and *persistent* income shocks. Suppose household i receive after-tax income Y_{it}, which takes the form:

$$\log(Y_{i,t}) = \kappa_t + y_{i,t},$$
$$y_{i,t} = P_{i,t} + \epsilon_{i,t},$$

where $P_{i,t} = \rho P_{i,t-1} + \zeta_{i,t}$, with $\rho < 1$ governing the persistence of earnings. Persistent shocks ζ are such that $\zeta_{i,t} \sim \mathcal{N}(0, \sigma_\zeta)$ and temporary shocks ϵ are such that $\epsilon_{i,t} \sim \mathcal{N}(0, \sigma_\epsilon)$. Additionally, the shocks $\zeta_{i,t}$ and $\epsilon_{i,t}$ are independent over time and across households.

Details and suggestions

- The PSID became bi-annual after 1997. Use the years 1978-1997. Assume we are in a steady state (i.e., just estimate a single variance for permanent and temporary income). To align with Blundell et al. (2008) and much of the literature that uses the PSID drop the individuals from the SEO oversample.

  Answer: See Part 1 in Appendix: Data code for details.

- You'll want to impose some sample selection criteria. Be clear on how you set your sample criteria. Often people only use observations above a minimum earnings cutoff and require that individuals satisfy that cutoff [X] times to be in the sample. Make a table with some descriptive statistics on your sample.

  Answer: See Part 2 in Appendix: Data code for our sample selection criteria. We are looking at individuals with ages between 30 and 65, that are married, and we dropped all individuals whose income is in the first and last 10% of household labor income distribution. Table 1 presents some statistics for income after taxes and transfers, years of education, number of members in the household, and annual worked hours.

*Email: fllopes@wisc.edu.
†Email: lordmedrano@wisc.edu.
‡Email: yyun23@wisc.edu.

Variable	Obs	Mean	Std. Dev.	Min	Max
Income	56,782	33,335.35	15,518.95	4,921.96	606,021.90
Years of education	56,395	13.1	2.31	1	17
Members in the household	56,782	3.53	1.25	1	12
Annual worked hours	55,843	1735.34	928.30	0	12,635

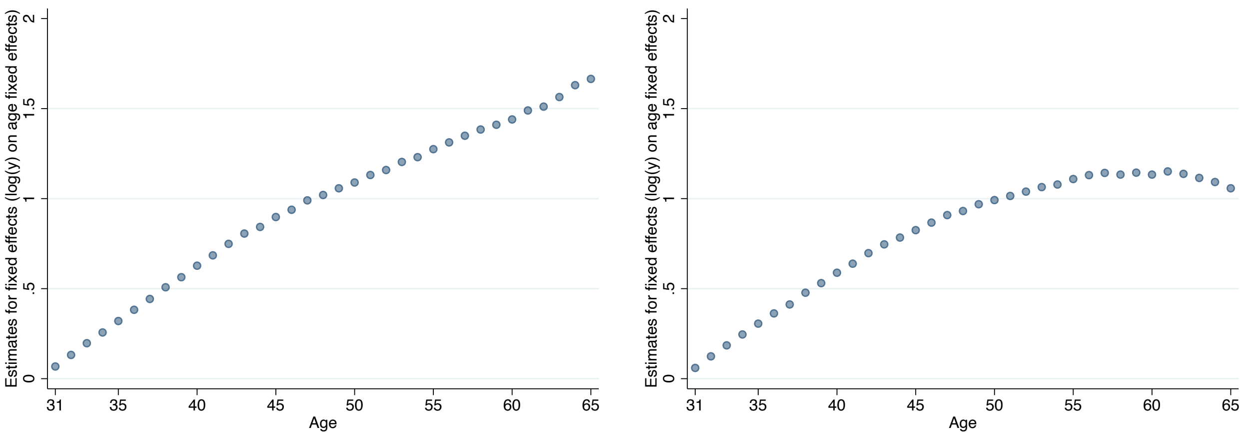
Table 1: Summary statistics

- Don't forget to remove the life-cycle component of earnings, κ_t. Be clear on how you choose to do it. Make a graph of your predicable age earnings profile κ_t.

  Answer: To compute the residual income, we compute the following panel regression for log income after taxes and transfers using age-fixed effects:

$$\log(Y_{i,t}) = \alpha + \sum_{s=30}^{65} \beta_s \times \mathbb{I}\{age_{i,t} = s\} + \theta_{i,t}.$$

  It is important to mention that we are using variable `i11113` from the PSID as our measure of income after taxes and transfers for the households[1]. Figure 1a display the estimated parameters for age. Alternatively, instead of using variable `i11113` for income, we use variable `i11103` that represents the household labor income. Figure 1b presents the estimates for the age fixed-effects. Note that we observed a hump-shaped curve for age-specific labor income, but the age-specific income after taxes and transfers are increasing in age.



(a) Post-govt income

(b) Labor income

Figure 1: Estimated coefficients

- To get started set $\rho = 0.97$. Make a table with your estimate of σ^2_ϵ and σ^2_ζ.

  Answer: Once we estimated the previous model, we compute the residuals (i.e., residual income):

$$\hat{\theta}_{i,t} = \log(Y_{i,t}) - \hat{\alpha} - \sum_{s=30}^{65} \hat{\beta}_s \times \mathbb{I}\{age_{i,t} = s\},$$

[1] This variable represents the combined income after taxes and government transfers of the head, partner, and other family members.

and then compute the "pseudo"-difference for the residual income given by

$$\tilde{\Delta}\hat{\theta}_{i,t} = \hat{\theta}_{i,t} - \rho\hat{\theta}_{i,t-1},$$

which allow us to compute the variance of the persistent and transitory shocks, respectively, as:

$$\sigma_\epsilon^2 = -\frac{1}{\rho}Cov\left(\tilde{\Delta}\hat{\theta}_{i,t}, \tilde{\Delta}\hat{\theta}_{i,t+1}\right) \text{ and } \sigma_\zeta^2 = \frac{1}{\rho}Cov\left(\tilde{\Delta}\hat{\theta}_{i,t}, \rho^2\tilde{\Delta}\hat{\theta}_{i,t-1} + \rho\tilde{\Delta}\hat{\theta}_{i,t} + \tilde{\Delta}\hat{\theta}_{i,t+1}\right).$$

Given the subsample that we are using, our estimates of the variance of the persistent and transitory shocks for both measures of income are presented on Table 2. For more details, see Part 4 in Appendix: Data code.

Shock	Variance	
	Post-govt income	Labor income
Persistent (σ_ζ^2)	0.0166	0.0272
Transitory (σ_ϵ^2)	0.0174	0.0210

Table 2: Variance of Income Shocks

Part 2: Model

Use your estimates on the labor income process from Part (1) in a life-cycle Bewley Model. Solve and simulate data from a model with the following features (essentially the model from Kaplan and Violante (2010) without retirement where the labor income process has persistent and transitory shocks):

- Agents enter the model with zero assets and are in the model for 35 years. There is no retirement.
- Agents labor income follows the income process from Part (1). Use your estimates from Part (1) for the variance of transitory and persistent shocks.
- Suppose initial permanent income $P_{i,0}$ is drawn from a normal distribution with mean zero and variance $\sigma_{\zeta 0}$. To get started use the variance from Kaplan and Violante (2010).
- Suppose agents receive utility from consumption and borrow/save at an exogenously given interest rate $r > 0$. To get started set $r = 4\%$ and the discount factor of agents $\beta = 0.975$. Set the borrowing constraint to either zero or the natural borrowing limit. Be clear on the choices that you make.

Write down a recursive representation of this model. Solve and simulate data from this model. To get started, simulate 1,000 agents for a full life cycle. Using the simulated data do the following:

1. Make a graph of the average value of wealth by age in your model.
2. Plot the variance of consumption (i.e., consumption inequality) in your model.

3. Using model simulated data, complete the insurance coefficients using the Blundell et al. (2008) methodology. How do your estimates compare to the estimates reported in their paper.

4. Finally, using model simulated data compute the true insurance coefficients (equation (4) in Kaplan and Violante (2010)). How do these coefficients compare to the results from the step above. Do your results align with Kaplan and Violante (2010)? If they do not align, what features of your model relative to Kaplan and Violante (2010) do you think led to this.

Answer: The state space of the household's problem can be defined as a 4-dimensional space with 4 state variables: current asset a, age-dependent income κ, permanent income z, and transitory income ϵ. The problem could be written in an recursive form as below. The second line follows from discretization of the state space $\{z, \epsilon\}$ following Tauchen (1986).

$$
\begin{aligned}
V(a, \kappa_t, z, \epsilon) &= \max_{\{c, a'\}} \; u(c) + \beta \mathbb{E}\left[V(a', \kappa_{t+1}, z', \epsilon')\right] \\
&= \max_{\{c, a'\}} \; u(c) + \beta \sum_{z', \epsilon'} \Pi(z', \epsilon' | z, \epsilon) V(a', \kappa_{t+1}, z', \epsilon') \\
V(a, \kappa_{T+1}, z, \epsilon) &= 0 \\
s.t. \quad & c + a' = (1+r)a + \exp(\kappa_t + z + \epsilon) \\
& a' \geq -\mathcal{A} \\
& 1 \leq t \leq T
\end{aligned}
$$

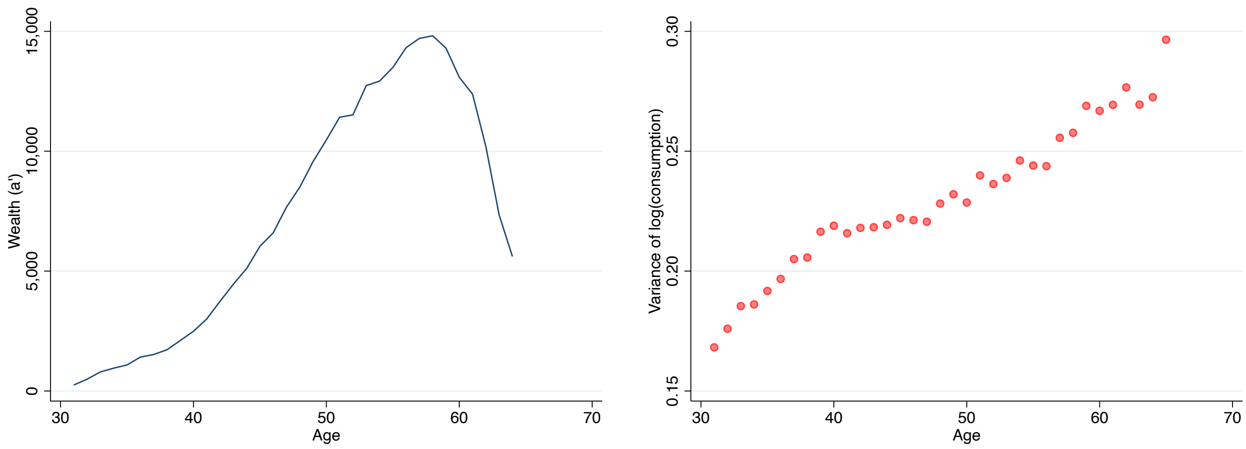
Note that the process of κ_t is deterministic and has been estimated in Part 1: Data.[2] Given the orthogonality of the persistence and transitory shocks, the transition probability $\Pi(z', \epsilon' | z, \epsilon)$ for the income process is simply constructed as a product of transition probabilities: $\Pi(z', \epsilon' | z, \epsilon) = \Pi(z'|z) \times \Pi(\epsilon'|\epsilon)$. For more details regarding the code, see the attached Julia code.

Technically, we impose 39 grid points and 19 grid points for z and ϵ to apply discretization. We also impose borrowing limit of zero, i.e. $\mathcal{A} = 0$, for simplicity and set the maximum possible value of a to be \$300,000 following the original setup in Kaplan and Violante (2010). In terms of obtaining the policy functions, we apply linear interpolation of the asset grid defined on $[0, 300000]$.

In Figure 2a we display the average wealth by year and variance of log-consumption for the simulated economy. We observe a similar average wealth curve as in Kaplan and Violante (2010) but with a different magnitude. Average wealth increases until around age 57 (i.e., $t = 27$) and begin to decrease after that as agents reduce savings as they tend to the end of their lives. On Figure 2b, we confirm that the consumption inequality measured by variance of consumption is increasing in age, or time from the beginning of the economy. Though the numbers are different, this also matches the data pattern in Blundell et al. (2008) qualitatively.

Additionally, our estimates of the insurance coefficients and the true insurance coefficients are displayed in Table 3. Qualitatively, we get similar results to Kaplan and Violante (2010) in the sense that households are better insuring themselves against transitory shocks, but, not that good for persistent shocks. In our model, insurance coefficients are calculated to be very low as compared to Kaplan and Violante (2010) when we look at both model-implied BPP estimates and true insurance coefficients. About 44% of transitory income shock is insured, which is less

[2] In this particular case, we have that the life-cycle component is given by $\kappa_t = \hat{\alpha} + \hat{\beta}_t$.



(a) Average wealth

(b) Variance of consumption

Figure 2: Wealth dynamics and consumption inequality

than half of the insurance coefficient reported by Kaplan and Violante (2010). For persistent income shock, only 2% of income shock is insured. This could be either because households die at retirement in this model, so that the income shocks in each period have larger effects on the overall lifetime income and has larger pass-through effects to consumption, or because of our technical coarseness in that we applied linear interpolation of asset grid points when obtaining the policy function for asset choices.

Insurance coefficients	Model BPP	Model True
Transitory $(1-\phi)$	0.4465	0.4458
Persistent $(1-\psi)$	0.0218	0.0238

Table 3: Insurance Coefficients

References

R. Blundell, L. Pistaferri, and I. Preston. Consumption Inequality and Partial Insurance. *American Economic Review*, 98(5):1887–1921, 2008.

G. Kaplan and G. L. Violante. How much Consumption Insurance beyond Self-Insurance? *American Economic Journal: Macroeconomics*, 2(4):53–87, 2010.

G. Tauchen. Finite State Markov-Chain Approximations to Univariate and Vector Autoregressions. *Economics Letters*, 20(2):177–181, 1986.

Appendix: Data code

```
/// Econ 810: Advanced Macroeconomic Theory
/// Professor: Carter Braxton
/// Problem Set 1: Variance of persistent and transitory shocks
/// Authors: Fernando de Lima Lopes, Stefano Lord-Medrano, Yeonggyu Yun
/// Date: 01/28/2023

///////////////////////////////////////////////////////////////////////////
* Housekeeping
* Clear workspace
clear all

* Set directory
cd "/Users/smlm/Desktop/Datasets - Metrics/PSID data"
use pequiv_long.dta
* Install coefplot package
ssc install coefplot, replace

///////////////////////////////////////////////////////////////////////////
* Part 1: Cleaning
* Drop observations after 1997 due to change in format of PSID
keep if inrange(year, 1978, 1997)
summ year

* Drop SEO oversample (see page 372 of Codebook for the Cross-National
* Equivalent File)
drop if x11104 == 12
summ x11104
///////////////////////////////////////////////////////////////////////////
* Part 2: Sample selection criteria
* Age between 30 and 65
keep if inrange(d11101, 30, 65)
* Marital status: married
keep if d11104 == 1
* Drop individuals that are in the first and last 10% of HH labor income distr.
keep if inrange(i11103, 11001, 84375)
* Describe data set and compute some statistics
describe
summarize
///////////////////////////////////////////////////////////////////////////
* Part 3: Remove the life-cycle component of earnings (\kappa_{it})
* Note: y_{it}=\log(Y_{it})-\kappa_{it}
* We choose age as our observable.
* Compute log HH post-government income (TAXSIM)
* See page 218 of Codebook for the Cross-National Equivalent File
* Drop negative values
drop if i11113<0
gen log_y = log(i11113)
* Define panel
xtset x11101LL year, yearly
* Run panel regression for log_y on age
xtreg log_y i.d11101, fe vce(robust)
* Plot coefficients of age on log-income
coefplot, vertical drop(_cons) noci nolabel
///////////////////////////////////////////////////////////////////////////
* Part 4: Compute variance of persistent (\zeta) and transitory (\varepsilon)
* shocks
* Compute residuals of panel regression and generate pseudo differences given by
* \Tilde{\Delta}y_{t}=y_{t}-\rho y_{t-1}
```

```
predict res_y, res
gen pseudo_dif_t0 = res_y - 0.97*L1.res_y
gen pseudo_dif_aux = (0.97^2)*L1.pseudo_dif_t0+0.97*pseudo_dif_t0+F1.
    pseudo_dif_t0
* Transitory shock variance
corr pseudo_dif_t0 F1.pseudo_dif_t0, cov
display (-1/0.97)*(-0.016923)
* Permanent shock variance
corr pseudo_dif_t0 pseudo_dif_aux, cov
display (1/0.97)*(0.016147)
```

Listing 1: Estimation of persistent and transitory shock variances (Stata)