U.S. Income Mobility for Men: 1979-2017

Soohyun Choi

University of Texas at Dallas

2022 Midwest Economics Association Annual Meeting

Contact: Soohyun.Choi@utdallas.edu

Four concepts of income mobility[1]

Intertemporal income mobility

1. Positional change
2. Individual income growth
3. Long-term inequality
4. **Income risk: Each person's period-specific income is the sum of a permanent component (the long-term average) and a transitory component (the period-specific deviations from the average)**
   - The transitory volatility is the source of income risk.

---

[1] Jäntti and Jenkins(2015)

Implication of variance decomposition

Increase in permanent variance

- Causes income distribution to widen over time
- Rankings are preserved
- Possible determinants: Labor demand shift, education

Increase in transitory variance

- Shuffles income rankings
- Implies higher income risk
- Possible determinants: Globalization, deregulation, de-unionization, and temporary employment

Data: CPS

The Annual Social and Economic Supplement (ASEC) of the Current Population Survey (CPS)

- The publicly-available version downloaded from the Center for Economic and Policy Research (CEPR)
- Ranges from 1979 to 2017
- Restrict to men between ages 30 and 59, who are not full-time students, with positive earned income and non-missing educational attainment information
- Drop zero-weighted samples
- Converted to 2017 CPI-U-RS dollars
- Trim the top 4% to eliminate top-coded incomes

Descriptive Statistics: CPS Cross-Section

	Mean	Standard Deviation	Minimum	Maximum
Age	43	8.424	30	59
Married (%)	0.74	0.437	0	1
Race:				
White (%)	0.78	0.415	0	1
Black (%)	0.08	0.275	0	1
Hispanic (%)	0.09	0.283	0	1
Others (%)	0.05	0.219	0	1
Education:				
Less than high school (%)	0.12	0.329	0	1
High school (%)	0.33	0.472	0	1
Some college (%)	0.25	0.432	0	1
College (%)	0.19	0.391	0	1
Advanced (%)	0.11	0.308	0	1
Employment:				
Full time, full year (%)	0.82	0.385	0	1
Working hours per week	43.56	9.46	1	99
Working weeks	48.82	8.87	1	52
Wage and Salary (2017 Dollars)	56,584	33,668	1	200,000

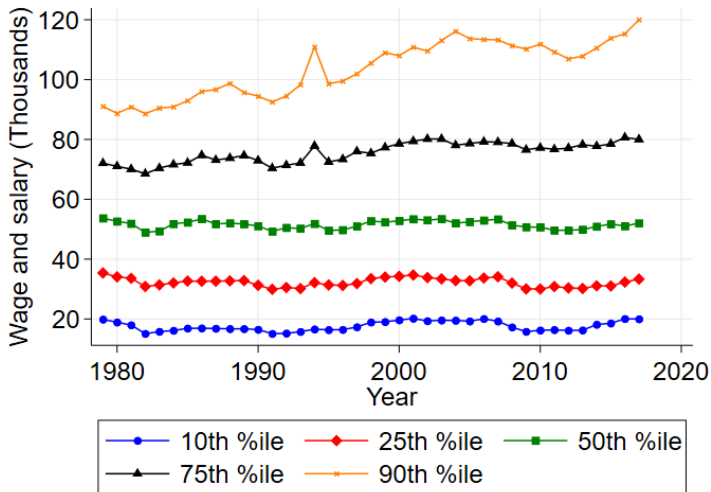# Data overview: CPS cross-section

Figure: Male earnings by percentiles

Data: Pseudo Panel

- In the CPS, individuals are followed at most two years.

- To investigate longer-term earnings patterns, a **pseudo panel** is constructed:
  Each individual is classified into only one cohort where the characteristics for creating cohorts are exogenous and time-invariant.

- Based on an individual's year of birth, education level, and race

Methodology: Obtain residuals

- Regress log earnings on education, an age polynomial, and interactions between age, education variables, separately by calendar year → Obtain residuals

$$z_{ct} = \beta_0 + X'_{ct}\beta_1 + Y'_{ct}\beta_2 + Z'_{ct}\beta_3 + \epsilon_{ct} \qquad (1)$$

where z_{ct} is log earnings for cohort c and time t, X_{ct} is a vector of five education dummy variables, Y_{ct} is an age polynomical (cubic), and Z_{ct} is interaction between education dummy and age.

- The regressions are weighted by the square root of the cohort size

Methodology: Extended Semiparametric (ESP) Model

- Developed by Moffitt and Zhang (2018)
- Overcomes one criticism on the widely used error component (EC) model, under which estimates are often sensitive to parametric assumption.

$$\epsilon_{cat} = \underbrace{\alpha_t \mu_{ca}}_{\text{Permanent Component}} + \underbrace{\beta_t \nu_{ca}}_{\text{Transitory Component}} \tag{2}$$

ϵ_{cat}: Log earnings residual for cohort c at age a and year t
α_t and β_t: Calendar time shifts

Methodology: Extended Semiparametric (ESP) Model

Permanent Component:

$$\mu_{ca} = \mu_{c0} + \sum_{s=1}^{a} \omega_{cs} \tag{3}$$

Transitory Component:

$$\nu_{ca} = \xi_{ca} + \sum_{s=1}^{a-1} \psi_{a,a-s} \xi_{c,a-s} \text{ for } a \geq 2 \tag{4}$$

$$\nu_{c1} = \xi_{c1} \text{ for } a = 1 \tag{5}$$

$|\psi_{a,a-s}| < 1$

ω_{cs}: A permanent shock

$\xi_{c,a-s}$: A transitory shock

Methodology: Extended Semiparametric (ESP) Model

- ω and ξ are nonparametric functions of age
- ψ are nonparametric functions of age and leg length

$$Var(\omega_{ca}) = e^{\sum \delta_j (a-25)^j} \tag{6}$$

$$Var(\xi_{ca}) = e^{\sum \gamma_j (a-25)^j} \text{ for } a \geq 2 \tag{7}$$

$$Var(\xi_{c1}) = k e^{\sum \gamma_j (1-25)^j} \text{ for } a = 1 \tag{8}$$

$$\psi_{a,a-b} = [1 - \pi(a-25)][\sum \omega_j e^{-\lambda_j b}] + \sum \eta_j D(b = j) \tag{9}$$

- The degree of the expansion is chosen by generalized cross-validation (GCV)

Methodology: Generalized method of moments (GMM)

- Using a generalized method of moments (GMM), the estimator finds close matches for population variances and autocovariances in equations (2)-(9) to their sample counterparts from log earning residuals $\hat{\epsilon_{ct}}$

- Since estimation with the weighing matrix can lead to biases in finite samples (Doris et al., 2011), an identity matrix can be chosen as an alternative (Altonji et al., 2013). $\rightarrow$ Minimum distance method

Result: Gross volatility

Gross volatility = The variance of first-differenced residuals
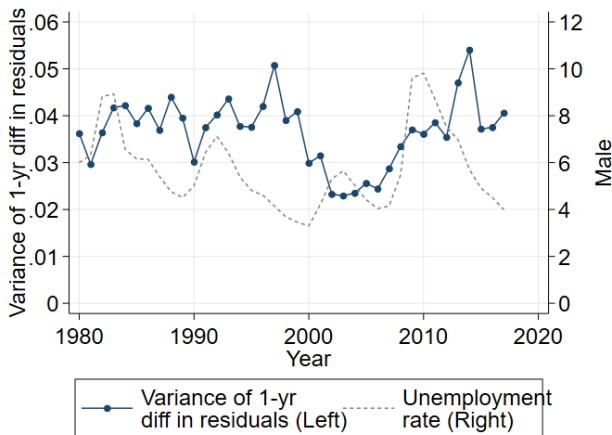

Figure: Gross volatility of male log earnings residuals

Note: The variance is a weighted sum by cohort size in CPS.
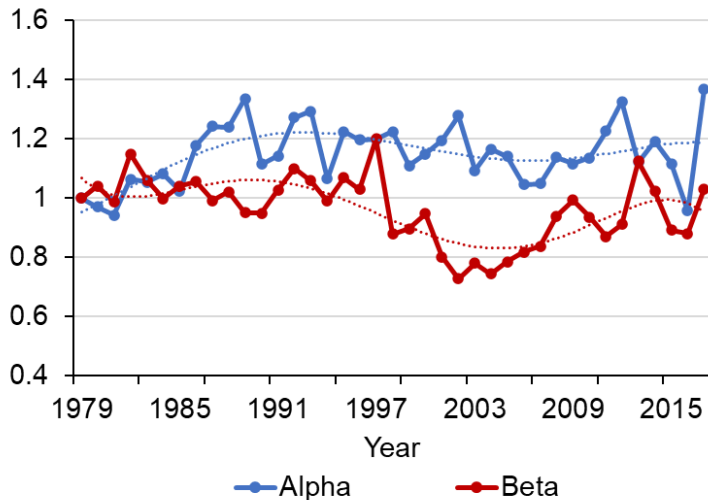
Result: ESP model

Figure: Extended semiparametric (ESP) model estimates of alpha and beta
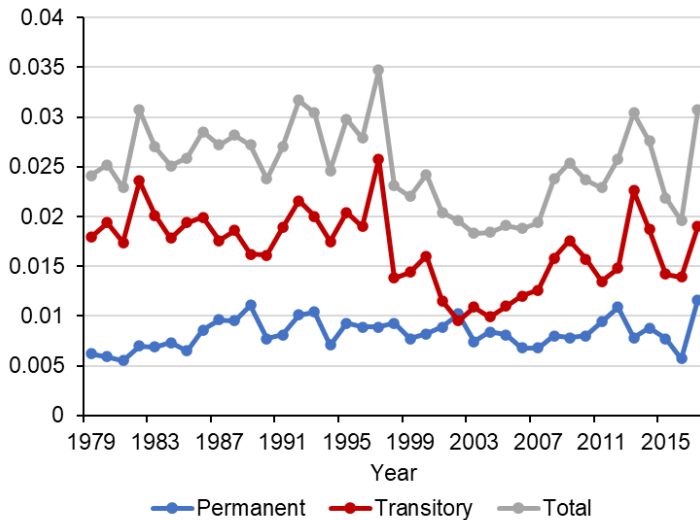
Result: ESP model

Figure: Fitted permanent, transitory, and total variance of log earnings residuals:

Conclusion

- Gross volatility: Consistent with the recent study (Moffitt et al., 2021) that shows little evidence of any significant trend in male earnings volatility since the mid-1980s.

- Income mobility: The transitory variance fluctuated through the mid-1990s and declined until 2002. Since then, the transitory variance increased through 2013 and almost recovered to the level in the mid-1990s.

- About 70% of the total variance can be explained by the transitory variance in average, which especially causes the increase in the total variance during Great Recession. (Countercyclical volatility pattern)

Further Research

- Use the restricted-use version
  - To protect the confidentiality of respondents, incomes in the CPS are top coded. The restrict-use version has higher top-coding thresholds.
  - The CPS uses hot-deck imputation for income item non-responses. The restricted-use version contains a flag variable of whether income is imputed or not.
  - The detailed geographical information at census-tract level
- Adopt the recent effort to reconcile the differences in gross volatility across studies (ex. The size of the left tail in income distribution)
- Estimate the EPS model at sub-group level (ex. Gender, Education)