The evolution of lifetime earnings in the US: evidence from Social Security earnings records
by
Michael V. Leonesio & Linda Del Bene

Discussant:
Stephen P. Jenkins
(ISER, University of Essex)
Background & Motivation

What’s been happening to lifetime earnings patterns in the USA over second half of 20th century?

• Conceptual: arguably lifetime earnings ‘a better indicator of well-being’ rather than short-run earnings

• Policy: in the USA, an individual’s Social Security retirement income (i.e. public pension income) is based on lifetime earnings

• There is a perceived increase in earnings uncertainty: are individual earnings becoming more variable?
Specific questions addressed

1. What is the trend in (cross-sectional) earnings inequality?
   • for men and women; for long period (1951–2001)
   • Do trends in lifetime inequality track those for short-run inequality?
   • Has the shape of lifetime age-earnings profiles changed over time?
   • Has the transitory component of earnings inequality increased relative to the permanent component?

Answers from US Social Security Administration’s Continuous Work History Sample (CWHS) data
The CWHS: advantages

‘Active file’: administrative records on earnings for about 3.1 million individuals = c. 1% all Social Security Numbers ever issued

Annual earnings data from 1951 onwards

- Very large sample size
- High degree of accuracy in reported earnings amounts
The CWHS: limitations

• Few covariates (e.g. date of birth, sex, race)

• **Coverage** of workers not universal
  – Wage and salary workers in ‘covered’ jobs (i.e. no self-employed)
    • Coverage rose from c. 61% in 1951 to c. 96% in 2001
  – Reasons for absence of recorded earnings not known
    • Zero earnings may mean not working or uncovered job
    • Low earnings in year may represent part-year working or other

• **Censoring**: only annual amounts up to maximum earnings liable for SS payroll tax recorded
  – Over 1951–2001, proportion attaining maximum declined from 25% to 5% (36% in 1965), but varies by sex:
    • 1951: 35% men censored, 3% women
  – Hence: values imputed for right censored observations
    • Assumed lognormal earnings distribution in each year
Sample selection criteria

- Cohorts born 1920–1966 ($n = 47$)
- Earnings at each age 31–61
  - Complete histories only for cohorts born 1920–40
- Drop those with any self-employment earnings
- Drop those ever received SS disability benefits
- Plus some other criteria in longitudinal analysis
  - Must have survived to 2001
  - Must have earned, or be on course to earn, enough to be eligible for retirement benefits at age 62
Measure of lifetime earnings (‘\textit{AIAE}’)

Cohorts born 1920–1940:
\begin{itemize}
  \item \textit{AIAE} = average between ages 31 and 61 of annual earnings at each age, indexed to age 61 using an average earnings index
\end{itemize}

Cohorts born after 1940:
\begin{itemize}
  \item \textit{AIAE} = average between ages 31 and age in 2001 at age \(a\), indexed to 2001 using an average earnings index
\end{itemize}
1. Cross-sectional earnings inequality (men)

- Only for 1979 onwards, because of high censoring prevalence before then
- Similar results if trim top 5% or top 10%
- Gini, percentile ratios
- Arguably similar to trends from CPS data (but 1990?)

% change in Gini value from 1979, men, uncensored
1. Cross-sectional earnings inequality (women)

- Full period 1951–1971, omitting 1963–5 (censoring greater than 10%)
2. Trends in lifetime earnings inequality

- 1920–1950 cohorts only (given incomplete histories)
- Striking rise in Gini for men, but not for women
- Argue that men’s result plausible despite high prevalence of censoring for early cohorts
3. Changing earnings profiles? (men)

Figure 4a. Lifetime earnings profiles, ages 31-61, 1920-66 birth cohorts

Median real earnings, men

Calculation of medians excludes zero earnings years.
3. Changing earnings profiles? (women)

Figure 4b. Lifetime earnings profiles, ages 31-61, 1920-66 birth cohorts

Median real earnings, women

Calculation of medians excludes zero earnings years.
4. Trends in components of variability

“Somewhat sceptical of result”: (1) differences from Moffitt & Gottschalk (rising transitory component). (2) Procedure attributes some transitory to permanent...
Comments (1)

• Very interesting paper
• Incredibly rich data set
  – Cf. UK DWP’s LLMDB and pension issues
• Would help to focus on more limited number of issues
  – E.g. why look at real earnings growth at different phases of lifecycle? Less about cross-sectional trends?
• Coverage issue of different nature for women compared to men?
  – Interpretation of differences between sexes, given secular increases in female participation rate?
  – [Graphs of this; trends in % covered, etc.]
Comments (2)

- Why restrict samples to those who survived to 2001?
- And yet also include some cohorts with incomplete histories
  - Differential mortality effects?
- Any effects associated with WWII?
- Censoring and results for men
  - Elaborate how used log earnings model to impute (more than a footnote!). E.g. did you fit lognormal distribution, with appropriate adjustment for right-censoring? Were covariates used? Etc.
  - Why log-normality?
    - Cf. GB(2) distribution fitted to top-coded CPS data by Burkhauser et al. (JBES, 2006);
    - Gini and other cross-sectional inequality estimates can be derived directly from fitted distributions
    - Difficult to extend this imputation approach to longitudinal setting!
      - Fit dynamics earnings models to the data (see below) … with right-censoring!
Comments (3)

Estimation of (trends in) transitory and permanent variance components of log earnings: \( y_{it} = \mu_i + \nu_{it} \)

- RE Tobit of log(earnings) on 4th order polynomial in age; permanent component estimated as 5-period moving average of residuals; transitory component estimated as difference between residual and permanent component
  - Problem: estimate of permanent component, \( V(\mu) \), contaminated by transitory variability if \( \nu_{it} \) is autocorrelated or follows MA process

- Apply Gottschalk & Moffitt (Econ. J., 2002) method?: \( V(\mu) \) is \( \text{Cov}(y_{it}, y_{it-k}) \) for \( k \) sufficiently large that transitory components uncorrelated, regardless of ARMA structure
  - US PSID secular rise in permanent variance; rise in transitory variance during 1980s, fell after 1991 \((k = 5)\)

- How to estimate ARMA model with censoring?