Life-cycle dynamics and the permanent/transitory decomposition of earnings inequality

Lorenzo Cappellari
Università Cattolica Milano

Winter School on Inequality and Social Welfare Theory,
Alba di Canazei, January 12th 2010

- Earnings dynamics analyses are an essential complement to studies of wage inequality
  - Positive
  - Policy

- Studies of mobility across the quantiles of the wage distribution
  - Assessement of low pay traps
  - Markov chain approach

- Studies of life-cycle earnings dynamics
  - Cross-sectional vs permanent inequality
  - Earnings instability

- This talk considers the latter approach

Plan of talk

1. Meaning of permanent/transitory inequality decomposition in the wage inequality literature

2. Models of earnings dynamics

3. GMM estimation

4. Empirical applications:
   1. Tests of wage theories
   2. Wage inequality decomposition
   3. Variance components and labour market institutions
   4. Intergenerational mobility

Permanent and transitory inequality

- The basic permanent/transitory income model (Friedman 1957) postulates that

$$w_{it} = \alpha_i + \varepsilon_{it};$$
$$E(\varepsilon_{it}) = E(\alpha_i \varepsilon_{it}) = 0$$

- Thence

$$Var(w_{it}) = Var(\alpha_i) + Var(\varepsilon_{it})$$
$$Cov(w_{it} \, w_{is}) = Var(\alpha_i)$$

- According to Permanent Income Hypothesis, consumption (welfare) inequality depends only on inequality of permanent incomes. Transitory shocks do not matter (as long as they can be smoothed away and agents are risk neutral)

- Permanent inequality is a measure of immobility. Mobility only through transitory fluctuations.

Permanent and transitory inequality

- The permanent/transitory income idea has been widely applied by labour economists for analysing wage inequality and its components

- Moffitt and Gottschalk (1995) started the literature.

- More elaborated models that we will consider in this lecture

- Interest spurred by major changes in the wage distribution in the US and elsewhere.

- Were these changes driven by permanent or transitory wage components?

- # Positive

    - Permanent inequality is due to changes in the returns to skills, fits explanations based on skill biased technical change

    - Transitory inequality ("earnings instability") consistent with increased turbulence in the labour market, declining institutions, employment instability

- # Policy

    - Permanent inequality implies long-term segmentation. Calls for interventions on skill endowment of the low paid (e.g. training programs)

    - Transitory inequality should be welfare neutral with perfect capital market and risk neutral agents

Models of earnings dynamics

- The Friedman's formulation is far too simplistic.
- Lacks three main ingredients:
    1. Permanent earnings (the returns to permanent skills) should vary over the life-cycle
    2. Transitory shocks can be serially correlated
    3. Calendar time effects.

- From now on: log-earnings deviations from some cross-sectional mean (could also be residuals from a first stage regression)

- We'll now go through these aspects using the following notation:

$$w_{it} = w^P_{it} + w^T_{it} ,$$

$$E(w^P_{it}) = E(w^T_{it}) = E(w^P_{it}\, w^T_{it}) = 0$$
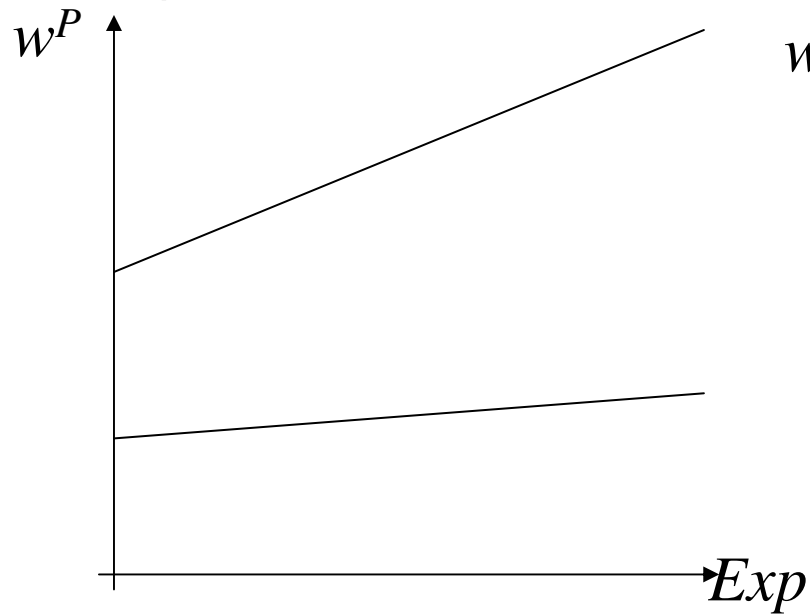
Life-cycle earnings

- Most popular specification is the *random growth* (RG) *model* (Hause, 1980; Baker 1997, Haider, 2001; Baker and Solon, 2003; Moffitt and Gottschalk 2008; Bingley et al 2009):
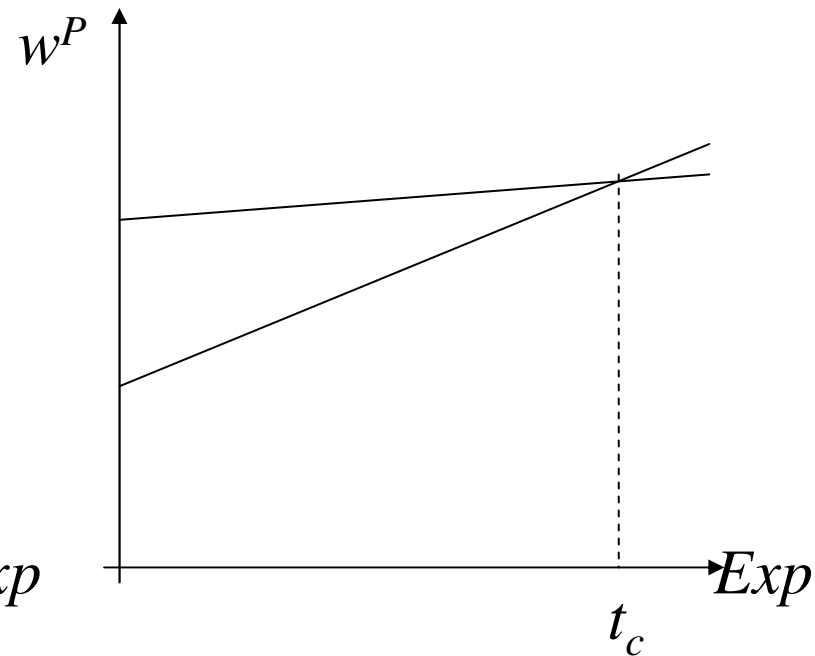
$$w^P_{it} = (\alpha_i + \beta_i EXP_{it});$$

$$(\alpha_i, \beta_i) \sim_i (0,0; \ \sigma^2_{\alpha}, \ \sigma^2_{\beta}, \ \sigma_{\alpha\beta})$$

Life-cycle earnings

$\sigma_{\alpha\beta}>0$ (signalling+matching; training and education complements)

$\sigma_{\alpha\beta}<0$ (on the job training); mobility also in permanent incomes

w^P

w^P

Exp

Exp

t_c

$t_c=-\sigma_{\alpha\beta}/\sigma^2_{\beta}$

9

- ## Alternatively (Dickens 2000, Cappellari and Leonardi 2010): *Random walk* (RW) *model*

$$w^P_{it}=r_{iexp(t)}=r_{iexp(t-1)}+\xi_{iexp(t)}$$

$$r_{i0}\sim_i(0, \sigma^2_r); \xi_{iexp(t)} \sim_{ie} (0, \sigma^2_\xi)$$

"Transitory" shocks

- Assume low order ARMA, e.g. ARMA(1,1).

- MaCurdy (1982):
  - Typical time series assumption that process is in steady state is untenable in this context (did not start in infinite past).
  - Non-stationarity won't cause estimation problems since consistency achieved over N, not over T.
  - Therefore, use time series process with an explicit treatment of initial conditions

$$w^T_{it} = v_{it} = \rho v_{it-1} + \varepsilon_{it} + \theta \varepsilon_{it-1};$$

$$\varepsilon_{it} \sim_{it} (0; \sigma^2_\varepsilon);$$
$$v_{i0} \sim_i (0; \sigma^2_0)$$

Time (and cohort) effects

- Modelling calendar time effects is crucial if we want to assess the contributions of variance components to changing inequality over time.

- In panel data, time effects are confounded with age effects.

- Solution: compute empirical earnings moments needed for estimation (see below) within narrowly defined birth cohorts, and stack them over cohorts, so that time and cohort effects can be separated.

- Then, model becomes

$$w_{it} = w^P_{it} + w^T_{it} =$$

$$= \pi_t \lambda_{c(i)}(\text{RG or RW}) + \tau_t \mu_{c(i)}(\text{ARMA}(1,1))$$

Minimum distance estimation

- (Chamberlain 1984; a nice exposition provided by Haider 2001)
- Application of GMM
- Derive "theoretical" covariance structure of earnings implied by the model, a function of model parameters
- Estimate parameters so as to minimise some distance between theoretical and empirical moments
- Empirical moments estimated by birth cohort, stacked over cohorts in estimation so as to separate time and cohort (age) effects.

"Theoretical" moments

- Permanent component with *RG*

$$Cov(w^P_{it} \, w^P_{is}) = [\sigma^2_\alpha + \sigma^2_\beta \, EXP_{it}EXP_{is} + \sigma_{\alpha\beta}(EXP_{it} + EXP_{is})] \, \pi_t\pi_s\lambda^2_{c(i)}$$

- Permanent component with *RW*

$$Cov(w^P_{it} \, w^P_{is}) = [\sigma^2_r + \sigma^2_\xi * min(EXP_{it}, \, EXP_{is})] \, \pi_t\pi_s\lambda^2_{c(i)}$$

Note different model implications: RG predicts wage inequality to grow over the life cycle at increasing rates (i.e. σ^2_β necessarily positive), whereas RW predicts linear growth.

- Transitory component with ARMA(1,1)

$$\text{Var}(w^T_{it}) = \mu^2_0 \lambda^2_{c(i)} \sigma^2_0 \qquad\qquad \text{if } t=0$$

$$\text{Var}(w^T_{it}) = \mu^2_t \lambda^2_{c(i)} [\sigma^2_\varepsilon (1+\theta^2+2\theta\rho)+\text{Var}(v_{it-1})\rho^2] \quad \text{if } t>0$$

$$\text{Cov}(w^T_{it} w^T_{its}) = \mu_t \mu_s \lambda^2_{c(i)} [\text{Var}(v_{is})\rho^2+\theta\, \sigma^2_\varepsilon] \text{ if } \qquad (t-s)=1$$

$$\text{Cov}(w^T_{it} w^T_{its}) = \mu_t \mu_s \lambda^2_{c(i)} [\text{Cov}(v_{it-1} v_{is})\rho] \text{ if } \qquad (t-s)>1$$

- Note: no close form solution but recursive structure due to the initial condition issue a-la-MaCurdy (1982)

Estimation

- For the time being assume a balanced panel of earnings data. Approach suited also for unbalanced panel (see below)

- $\Omega(\theta)$ "theoretical" covariance structure. A function of unknown parameters and –sometimes– conditional on empirical moments of observables, e.g. labour market experience

- Orthogonality assumption $E(w^P_{it}\, w^T_{it})=0$ implies that $\Omega(\theta)$ results from the sum of thoeretical moments of permanent and transitory earnings derived in previous slides.

- M_i empirical earnings moments for individual i, the cross-products of individual wages over time periods. $M=(1/N)\Sigma_i M_i$ sample moments

- $m_i=vech(M_i); \; \omega(\theta) = vech(\Omega(\theta))$

- The parameter vector is identified by the following set of moment restrictions:

$$E[m_i - \omega(\theta)] = 0$$

- A consistent estimate of θ can be obtained from the empirical counterpart of the moment restrictions, i.e. by minimising the distance between empirical and "theoretical" moments:

$$\theta_{MD} = \text{argmin}[m - \omega(\theta)]'W[m - \omega(\theta)]$$

where W is some suitable weighting matrix.

Estimation

- Efficiency would require $W = \text{var}(m)^{-1}$ (Optimally Weighted Minimum Distance, OWMD)

- However, correlation between sampling errors in second and fourth earnings moments may bias parameter estimates

- Altonji and Segall (1996) suggest to overcome the issue by setting $W=I$ and to reduce the variance of the estimator post-estimation using

$$V_{MD} = (G'G)^{-1}G'\text{var}(m)G(G'G)^{-1}$$

Where G is the gradient of $\omega(\theta)$ evaluated at the solution of the minimisation problem θ^*_{MD}

(Equally Weighted Minimum Distance, EWMD)

- The minimum distance estimator (a GMM estimator) is consistent, efficient and asymptotically normal:

$$\sqrt{N}\,(\theta_{MD}-\theta) \sim \aleph\,(0,V_{MD})$$

- Model specification can be tested against the alternative of unrestricted covariance structure using

$$(m-\omega(\theta^*_{MD}))'R^{-1}((m-\omega(\theta^*_{MD})) \sim \chi^2_{(q-p)}$$

$R=Avar(m)A$, $A=I - G(G'G)^{-1}G'$, q is the number of restrisctions and p the number of parameters to be estimated.

- Works on balanced and unbalanced panels.
- For a formal discussion of the unbalanced panel case see Haider (2001)
- Here I provide a practitioners' discussion
- The key is the computation of empirical moments

- In the balanced panel case, the vector of observations on individual earnings is

$$w_i' = (w_{i1}\ w_{i2}\ ....w_{iT})$$

so that $M_i = w_i w_i' =$

$$\begin{vmatrix} w_{i1}^2 & w_{i1}w_{i2} & .... & w_{i1}w_{iT} \\ w_{i1}w_{i2} & w_{i2}^2 & ... & w_{i2}w_{iT} \\ ... & ... & ... & ... \\ w_{i1}w_{iT} & w_{i2}w_{iT} & ... & w_{iT}^2 \end{vmatrix}$$

and M is the sample average of the M_i's

- Consider the unbalanced panel case

- Assume individual i lacks wage observation in period 2 (generalizes to any pattern of unbalancedness)

$$w_i' = (w_{i1} \ X \ ....w_{iT})$$

- Fill-in the missing value with a 0

$$w^*_i{}' = (w_{i1} \ 0 \ ....w_{iT})$$

- Derive unbalanced moments

$$M^*_i = w^*_i w^{*\prime}_i = \begin{bmatrix} w_{i1}^2 & 0 & \ldots & w_{i1}w_{iT} \\ 0 & 0 & \ldots & 0 \\ \ldots & \ldots & \ldots & \ldots \\ w_{i1}w_{iT} & 0 & \ldots & w_{iT}^2 \end{bmatrix}$$

- Define d_i' as the vector of dummies indicating presence in the panel, $d_i' = (1\ 0\ \ldots 1)$

- Matrix $D = \Sigma_i d_i d_i'$ provides the number of individuals contributing to earnings moments for any pair of time periods

- Then, empirical moments for the unbalanced panel case are given by matrix M^*, such as $cell_{kj}(M^*) = cell_{kj}(\Sigma_i M^*_i)/cell_{kj}(D)$, $cell_{kj}()$ denoting the cell of position kj

Data requirements

- Methods/models well suited for data sets that have large numbers of observations and few covariates

- Applications typically look at prime age men

- Good on administrative archives/register data

- Survey data could be problematic especially if one wants to define narrow birth cohorts

Applications

1. Tests of wage theories
2. Wage inequality decomposition
3. Variance components and labour market institutions
4. Intergenerational income mobility

Testing the on-the-job training hypothesis

- Need RG model
- The hypothesis implies $\sigma_{\alpha\beta} < 0$
- Hause 1980, Baker 1997, Haider 2001, Baker and Solon 2003, Bingley et al 2009, all find support for the theory
- Cappellari 2004 finds opposite results

# Testing the on-the-job training hypothesis

$t_c = 4.25$

Table 4: Core parameter estimates for the earnings model

	Baseline					
	Permanent component			Transitory component		
	Coeff.	S.E.			Coeff.	S.E.
σ^2_{α}	0.0079	0.0007		σ^2_{ε}	0.0627	0.0006
σ^2_{β}	0.0008	0.0000		σ^2_0	0.0337	0.0004
$\sigma_{\alpha\beta}$	-0.0034	0.0001		ρ	0.7732	0.0008

Bingley et al 2009, Danish data

Wage inequality decomposition

Figure 3: Fitted Permanent, Transitory, and Total Variances of Male Log Annual Earnings, Age 30-39
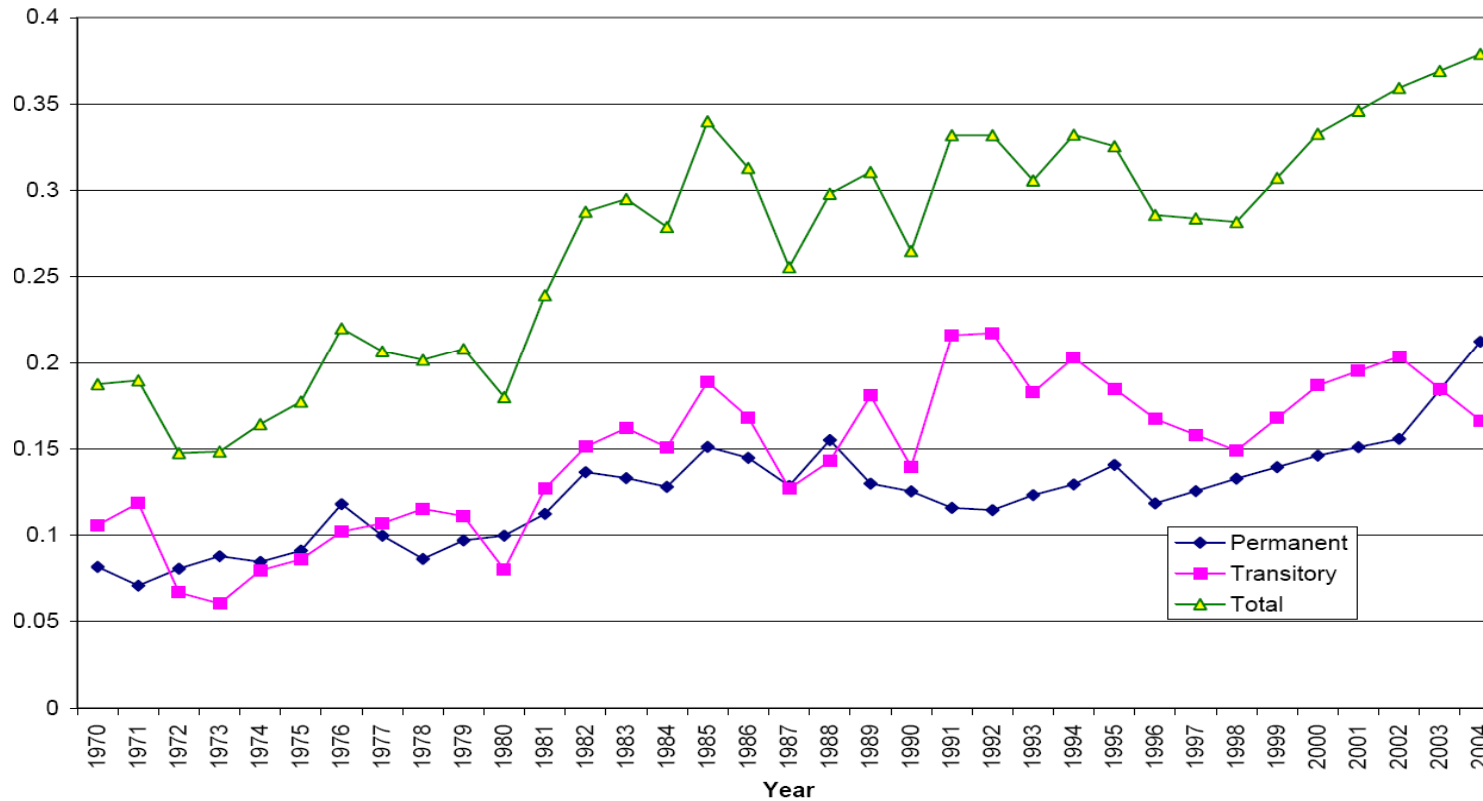
Moffitt and Gottschalk, 2008, US data

Wage inequality decomposition

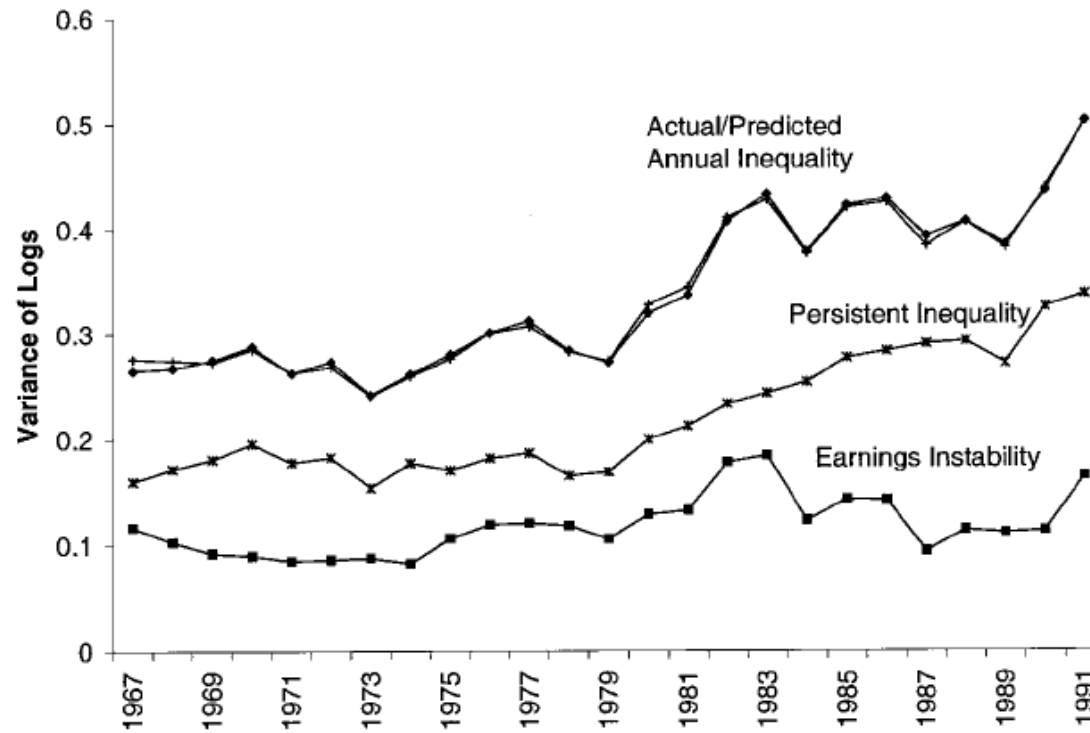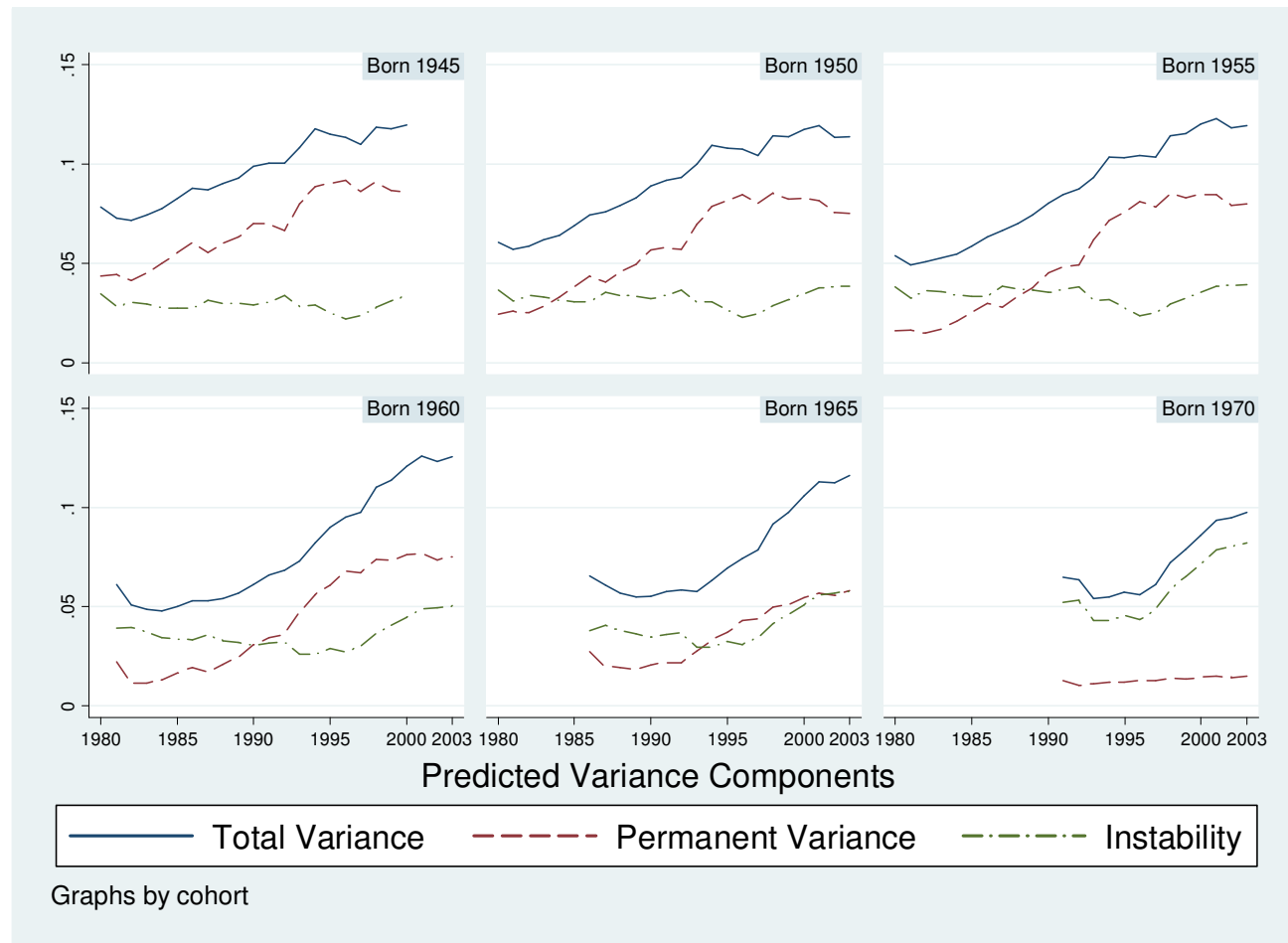824   Haider

FIG. 5.—Decomposition of annual earnings inequality with a random walk model. Data for this figure come from estimating the random walk model in equation (9).

Haider, 2008, US data

Wage inequality decomposition

Figure 2: Predicted variance components

Bingley et al 2009, Danish data

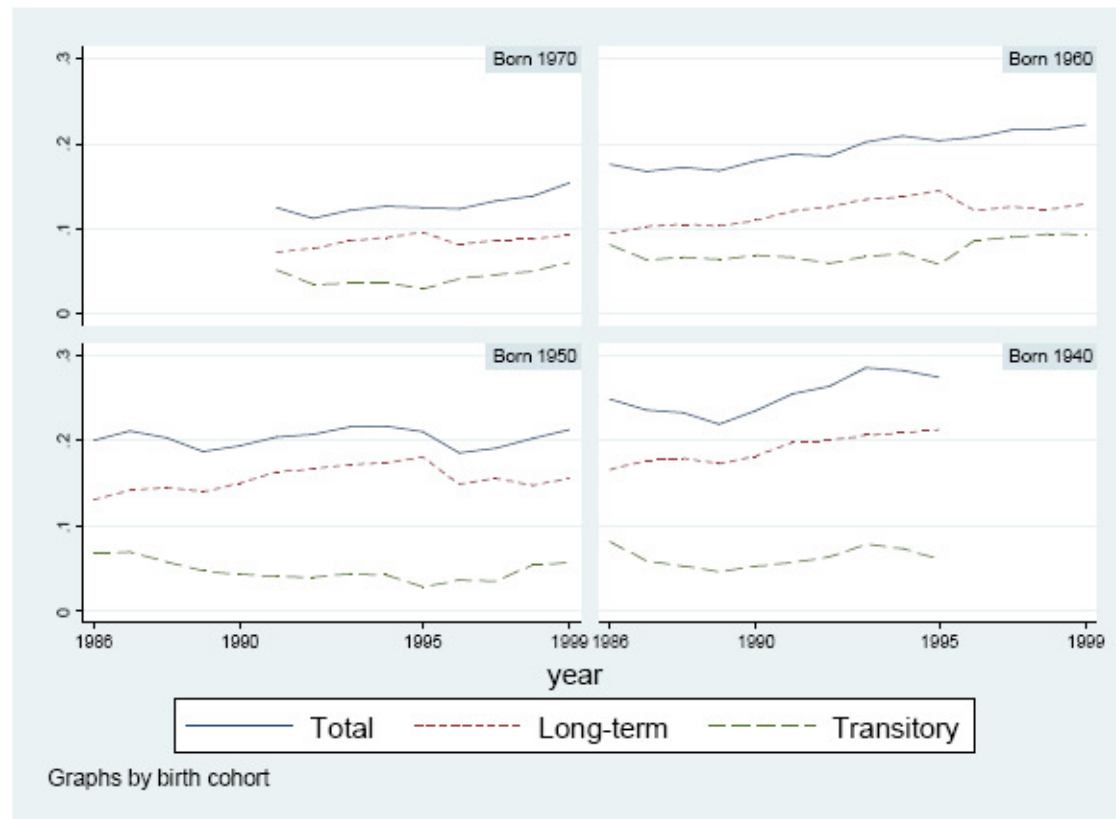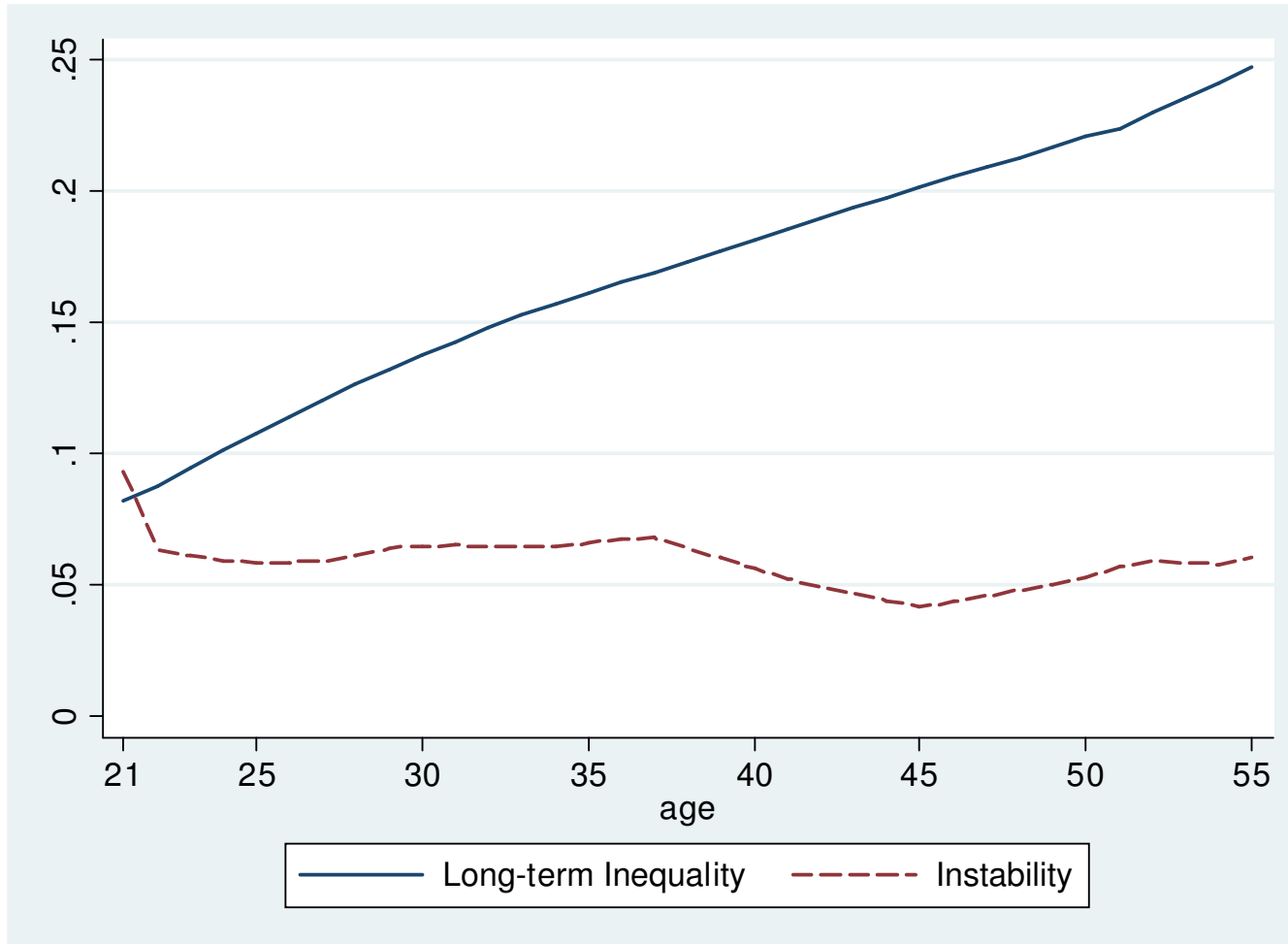# Wage inequality decomposition
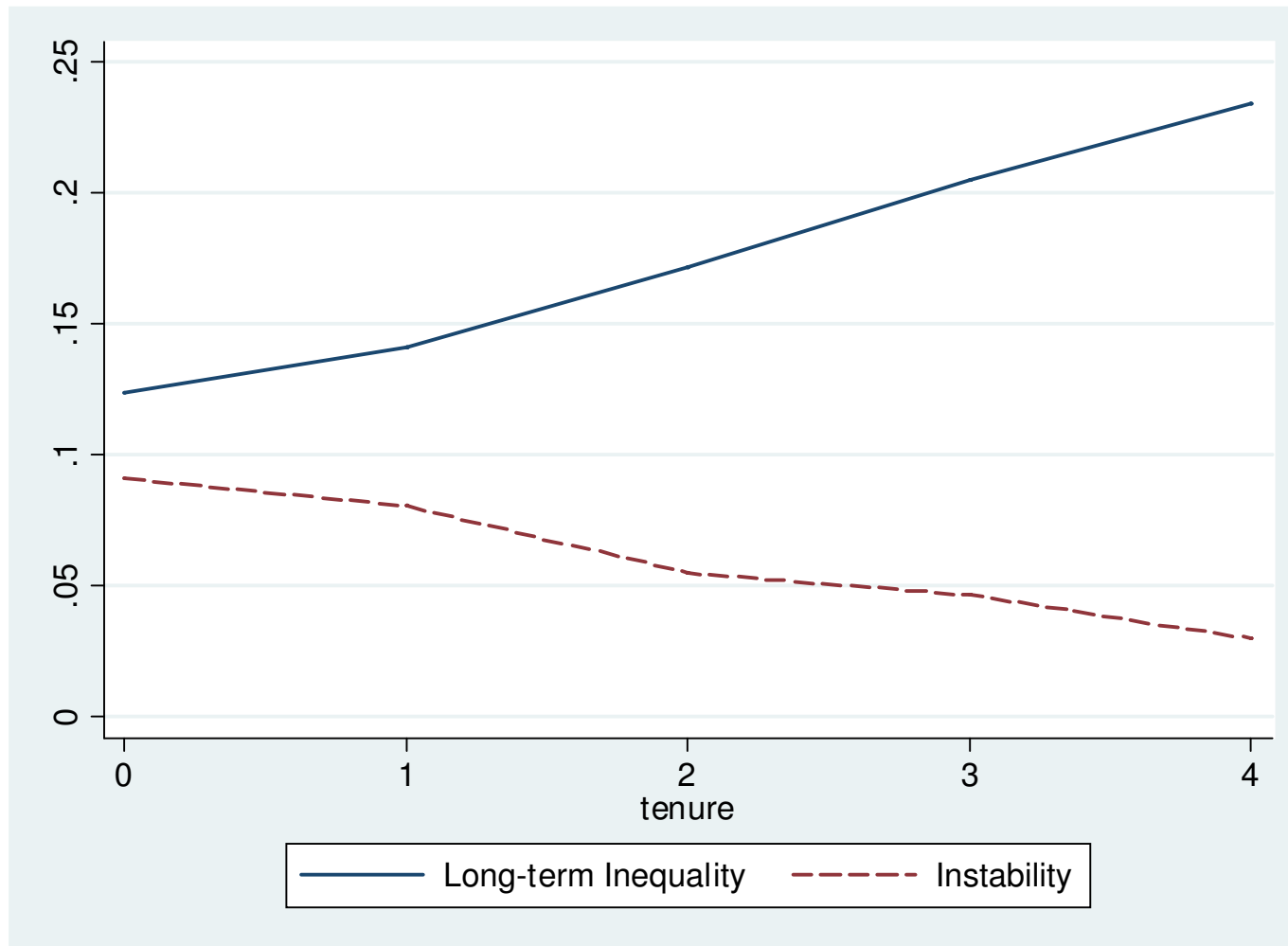

Figure 2: Predicted Variance Components

Cappellari and Leonardi, 2010, Italian data

Wage inequality decomposition

Cappellari and Leonardi, 2010, Italian data

Wage inequality decomposition

Cappellari and Leonardi, 2010, Italian data

Variance components and labour market institutions

- Earnings instability seems to play a role in explaining overall trends, and its relevance increasing in recent years
- What are its determinants?
- Labour market turbulence
- Relate variance components to declining labour market institutions that traditionally isolate the individual wage from turbulence
- Warning: models are formally identified (moment restrictions) but still evidence better interpreted as descriptive.

- We have already seen that on-the-job tenure is associated with lower instability
- Matching (Altonji and Pierret 2001; Lange 2007)
- Firm provided insurance (Guiso et al 2005)
- In a world of increased employment instability (e.g. the diffusion of temporary employment contracts) we should expect larger instability.

Variance components and labour market institutions

Table 5: Predicted Earnings Instability by Contract Type.

	Cohort born 1940	Cohort born 1950	Cohort born 1960	Cohort born 1970
Year 1988				
permanent contract	0.078	0.089	0.070	-
temporary contract	0.081	0.094	0.106	-
Year 1998				
permanent contract	-	0.058	0.082	0.052
temporary contract	-	0.086	0.122	0.081

Cappellari and Leonardi, 2010, Italian data

Variance components and labour market institutions

- Institutional arrangements may affect incentives
- Much is known about the disincentive effects that unemployment insurance induce in job search of the unemployed
- Are there effects also on employed individuals?
- Bingley et al (2009) look at this issue studying the wage effects of membership of unemployment insurance schemes in Denmark
- Use RG model + AR(1) instability, both as function of UI membership

Variance components and labour market institutions

- Extended model: Permanent wage

$$w^P_{ict} = \lambda_c \pi_t (\alpha_i + \beta_i EXP_{it} + \gamma_i F_{it} + \delta_i F_{it} EXP_{it});$$

$$(\alpha_i, \beta_i \gamma_i \delta_i) \sim [(0,0); (\sigma^2_\alpha \sigma^2_\beta \sigma^2_\gamma \sigma^2_\delta \sigma_{\alpha\beta} \sigma_{\alpha\gamma} \sigma_{\beta\delta})]$$

$$Cov(w^P_{ict} w^P_{ics}) = [\sigma^2_\alpha + \sigma^2_\beta EXP_{it} EXP_{is} +$$
$$\sigma_{\alpha\beta}(EXP_{it} + EXP_{is}) + \sigma^2_\gamma F_{it} F_{is} +$$
$$\sigma^2_\delta F_{it} EXP_{it} F_{is} EXP_{is} + \sigma_{\alpha\gamma}(F_{it} + F_{is}) +$$
$$\sigma_{\beta\delta}(F_{it} EXP_{it} + F_{is} EXP_{is})] \pi_t \pi_s \lambda^2_c$$

Extended model: Instability

$$\sigma^2_{\varepsilon ct} = \sigma^2_{\varepsilon}\, exp(\psi F_{ct})$$

Variance components and labour market institutions

- Find that when workers are covered by UI:

    1. Workers at the top (bottom) of the initial earnings distribution are shifted downward (upward)

    2. High growth workers slow down relative to lower growth ones

    3. Earnings instability (=shocks volatility) is higher

Variance components and labour market institutions

41
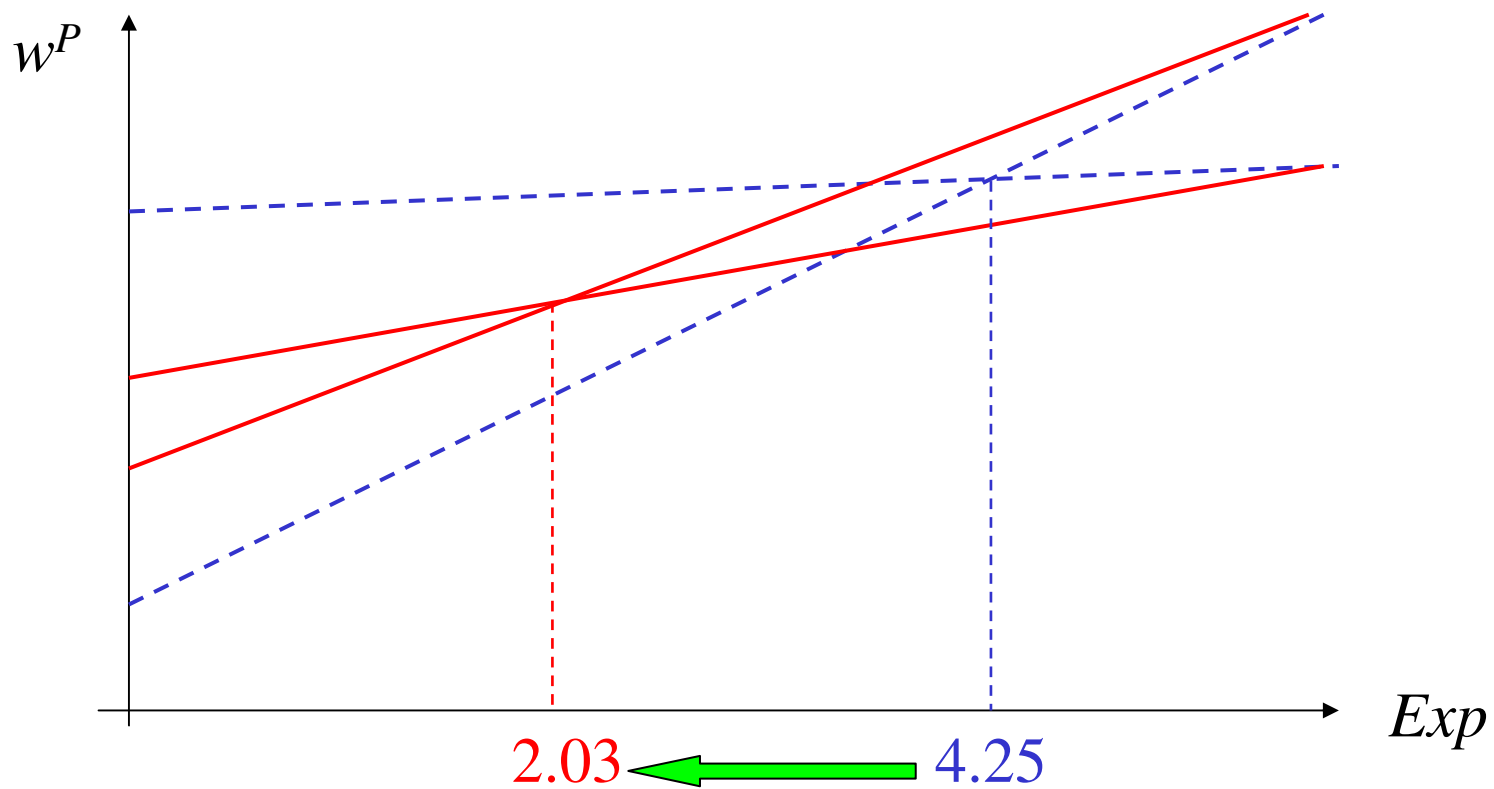
- **Interpretation**

  1. Changes in the distribution of entry wages are consistent with a reduced signalling value of education.
  2. Changes in the distribution of growth rates consistent with moral hazard: once insured lose incentives to accumulate skills
  3. Changes in instability consistent with moral hazard: e.g. once insured lose incentives to avoid shirking

- **Recall the initial words of caution.**

- **These are possible interpretations. Other based on self-selection are consistent with data, although robustness checks do not support them.**

Intergenerational earnings mobility (Björklund, Jäntti and Lindquist, JPubEc 09)

- Use register Swedish register data that allow following siblings income over time
- Wage (residual) for brother j of family i in year t

$$\varepsilon_{ijt} = a_i + b_{ij} + v_{ijt},$$

where a_i is a permanent component common to all siblings

b_{ij} is a permanent component unique to individual j,

v_{ijt} picks up deviations of annual from long-run income.

Intergenerational earnings mobility

- The first two components represent long run income, while the third is transitory income.

- A long-standing problem in this literature is to estimate intergenerational correlations in long run income, but typically researchers only observe total income (that also include income instability) which leads to downward bias.

- The BJL approach solves the problem

Given their assumptions

$$E\left[\varepsilon_{ijt}\varepsilon_{kls}\right] = \begin{cases} \sigma_a^2 + \sigma_b^2 + \sigma_v^2, & i = k; j = l; t = s \\ \sigma_a^2 + \sigma_b^2 + \lambda^{|t-s|}\sigma_v^2, & i = k; j = l; t \neq s \\ \sigma_a^2, & i = k; j \neq l; \forall t, s \\ 0, & i \neq k; j \neq l; \forall t, s \end{cases}.$$

So that

$$\rho = \frac{\sigma_a^2}{\sigma_a^2 + \sigma_b^2}.$$

- Note: in estimation, w_i becomes the vector of observation of income for one brother stacked over the vector of the other brother

- Can also use singletons, e.g. in a way similar to the unbalanced panel treatment we discussed earlier.
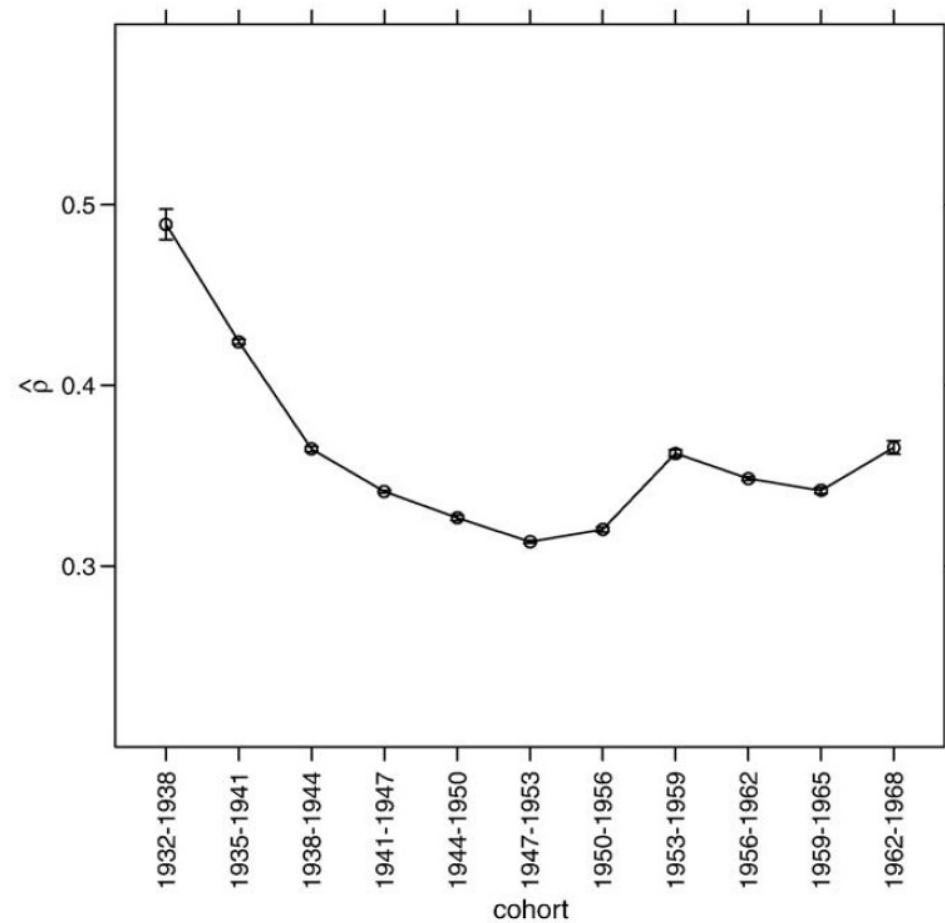
Intergenerational earnings mobility

Fig. 1. The estimated sibling correlation — baseline case. Note: First-stage regression includes a 3rd degree polynomial in age and year effects. The variance components are estimated using GMM on the brother (and singleton) over-time covariance matrices of residuals from first-stage regression. The baseline assumes that the within individual errors follow an AR(1) process.

47

Intergenerational mobility and life-cycle earnings

- Bingley and Cappellari (in progress)

- Do intergenerational transmission occurs via earnings growth?

- Extend the BJL model using the random growth specification:

$$w^P_{jit} = (\theta_j + \alpha_{ji}) + (\gamma_j + \beta_{ji})EXP_{jit}$$

- We identify new parameters of interest:

$$\sigma^2_\theta \; ; \; \sigma^2_\gamma \; ; \; \sigma_{\theta\gamma}$$

Conclusion

- Models/methods provide a way to assess the sources of changes in wage structure

- Much has been done on the wage distribution
  - Nice to have papers doing meaningful cross-country comparisons

- Still little work done on household incomes
  - Issues of aggregation over household members

- Fruitful extensions by using covariates in the model (factors associated with variance components)

- New fields of application: intergenerational correlations.

Reading list

- Baker M and Solon G (2003) Earnings Dynamics and Inequality among Canadian Men, 1976–1992: Evidence from Longitudinal Income Tax Records Journal of Labor Economics 21.
- Bingley, P., Cappellari, L. and Westergaard-Nielsen, N (2009): Flexicurity and wage dynamics over the life-cycle, mimeo.
- Björklund A, Jäntti M, Lindquist MJ (2009): Family Background and Income during the Rise of the Welfare State: Brother Correlations in Income for Swedish Men Born 1932-1968, Journal of Public Economics,
- Cappellari, L. (2004) The Dynamics and Inequality of Italian Men's Earnings: Long-term Changes or Transitory Fluctuations? Journal of Human Resources, 2004, vol. 39
- Cappellari, L. and Leonardi M. (2010) Earnings Instability and Tenure
- Chamberlain, G. (1984) Panel Data, in Handbook of Econometrics, vol. 2, edited by Griliches Z and Intriligator M (eds.) Amsterdam: North-Holland
- Dickens, R. The Evolution of Individual Male Earnings in Great Britain: 1975-95 Economic Journal, 2000, 110, (460), 27-49
- Gottschalk P. and Moffitt R. (2008) Trends in the Transitory Variance of Male Earnings in the U.S., 1970-2004, Boston College working paper 697.
- Gottschalk P. and Moffitt R. (2009) The Rising Instability of U.S. Earnings, Journal of Economic Perspectives 23
- Haider SJ (2001) Earnings Instability and Earnings Inequality of Males in the United States: 1967–1991 Journal of Labor Economics 19
- Hause, J. C. (1980) The Fine Structure of Earnings and the On-the-Job Training Hypothesis. Econometrica 48