Understanding Earnings Dynamics: Identifying and Estimating the Changing Roles of Unobserved Ability, Permanent and Transitory Shocks

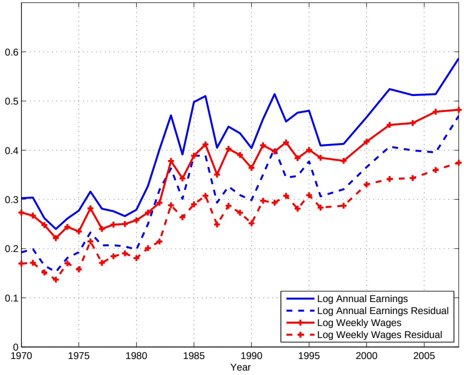
Lance Lochner (University of Western Ontario)
Youngki Shin (University of Western Ontario)

June 2014

Role of Unobserved Ability/Skills

- There is considerable interest in the evolution of inequality and the returns to ability/skill over time
- Widespread agreement that returns to observed skills (education, experience) have risen since the early 1980s
- Less agreement on role of unobserved skills
  - Autor, Katz and Kearney (2008) vs. Card & DiNardo (2002), Lemieux (2006)
- More generally, there is interest in understanding the factors driving the evolution of residual inequality

Earnings and Weekly Wage Inequality in the US

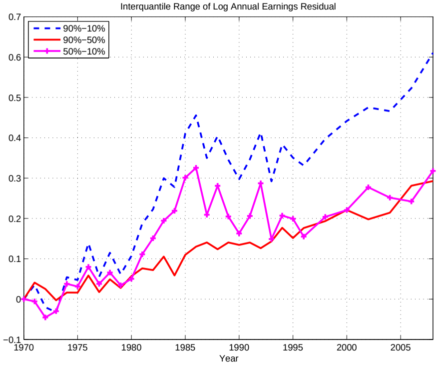


Source: 1970-2008 PSID

Residuals and Unobserved Ability/Skill

- CPS-based literature interprets all changes in residual inequality as changes in the 'pricing' of unobserved skills
  - e.g., Katz & Murphy (1992), Juhn, Murphy & Pierce (1993), Autor, Katz & Kearney (2008)
  - increased residual inequality reflects an increase in the 'returns' to unobserved skill
- Along with increase in returns to observed skill, this literature has motivated theories of SBTC (e.g. Acemoglu 1999, Caselli 1999, Galor & Moav 2000, Violante 2002)
- Changes in institutional factors and minimum wages may also be important (Card & DiNardo 2002, Lemieux 2006)
- Most recently, Acemoglu & Autor (2011) and Autor & Dorn (2012) argue that mechanization of routine tasks has led to polarization of skill demand

Interquantile Comparisons for Log Earnings Residuals, 1970-2008 PSID



What about Idiosyncratic Shocks?

- CPS-based literature largely ignores parallel research on earnings dynamics using PSID
  - labor: Gottschalk & Moffitt (1994, 2002, 2009, 2012), Haider (2001), Meghir & Pistaferri (2004), Robin & Bonhomme (2010)
  - macro: Heathcote, Perri & Violante (2010), Heathcote, Storesletten & Violante (2010)
- Decomposes variance of log wages/earnings residuals into permanent and transitory shocks over time
- Important for understanding consumption and savings behavior/inequality
- Estimates suggest similar increases in variance of both permanent and transitory shocks
- Transitory component unlikely to be related to unobs. skill
- Rarely account for changes in pricing of unobs. skills

Our Goal: Incorporating All Three Components

We consider a general log earnings/wage residual decomposition:

$$
\begin{aligned}
W_{it} &= \mu_t(\theta_i) + \kappa_{i,t} + \nu_{i,t} \\
\kappa_{i,t} &= \kappa_{i,t-1} + \eta_{i,t} \\
\nu_{i,t} &= \xi_{i,t} + \beta_{1t}\xi_{i,t-1} + \beta_{2t}\xi_{i,t-2} + \ldots + \beta_{qt}\xi_{i,t-q}
\end{aligned}
$$

- 'Unobserved Ability' literature effectively ignores any changes in distributions of $\kappa_{i,t}$ and $\nu_{i,t}$
- 'Earnings Dynamics' literature effectively ignores $\mu_t(\theta_i)$ or assumes $\mu_t(\theta_i)$ is time invariant

Earnings Components

$$
\begin{aligned}
W_{it} &= \mu_t(\theta_i) + \kappa_{i,t} + \nu_{i,t} \\
\kappa_{i,t} &= \kappa_{i,t-1} + \eta_{i,t} \\
\nu_{i,t} &= \xi_{i,t} + \beta_{1t}\xi_{i,t-1} + \beta_{2t}\xi_{i,t-2} + \ldots + \beta_{qt}\xi_{i,t-q}
\end{aligned}
$$

- $\mu_t(\cdot)$ reflects the pricing of unobserved skills θ, which may change over time due to technological change or institutional factors (e.g. unions, minimum wage)
- η_t reflects permanent idiosyncratic shocks like job displacement, switching employers, disability
- ν_t reflects potentially persistent but transitory shocks like temporary illness, family disruption, temporary demand shocks for employers

Changing Skill Prices vs. Shocks

- Changing skill prices affect earnings of similar individuals in the same way – strong co-movements over time
- Idiosyncratic shocks can move the wages/earnings of similar workers in very different directions
- We think of skill pricing functions as relatively slow moving based on supply/demand and institutional factors (e.g. unions, minimum wages)
  - likely to be more predictable
- Predictable changes in skill prices should have weak effects on within-cohort consumption inequality but should increase inequality across cohorts
- Increased variance of permanent shocks should increase within- and across-cohort consumption inequality
- Skill prices and variability of shocks have different implications for precautionary savings

What We Do – Outline

- We first consider conditions required for nonparametric identification
- Consider a moment-based approach for estimation
  - briefly discuss necessary conditions for identification with polynomial $\mu_t(\cdot)$ functions
  - provide Minimum Distance estimates assuming $\mu_t(\cdot)$ are linear/cubic polynomials
  - focus on log earnings residuals for men in PSID, 1970-2008

Nonparametric Identification: Simple Case

Consider non-parametric identification, beginning with a simple instructive case:

$$W_{it} = \mu_t(\theta_i) + \varepsilon_{it}$$

- ε_{it} are independent over time
- $\mu_t(\cdot)$ are strictly increasing

Some normalizations:

- $E(\theta) = E(\varepsilon_t) = 0$
- $\mu_1(\theta) = \theta$

Problem is very similar to that of measurement error literature (e.g. Hu and Schennach 2008)

Some Intuition on Identifying $\mu_t(\theta)$ and $f_\theta(\cdot)$

- If $\mu_t(\theta) = m_{0,t} + m_{1,t}\theta$, then the problem is just like a standard measurement error problem with multiple measurements
- If $W_1 = \theta$, we could just regress W_t on W_1 to identify $\mu_t(\cdot)$
- Due to ε_1, we would get attenuation bias
- Can use other $W_{t'}$ as instruments in a regression of W_t on W_1 to identify $\mu_t(\cdot)$
- Can then identify σ_θ^2
- General case is a bit like nonparametric IV in context of measurement error

Assumption 1

The following conditions hold for $T = 3$:

(i) The joint density of θ, W_1, W_2, and W_3 is bounded and continuous, and so are all their marginal and conditional densities.

(ii) W_1, W_2, and W_3 are mutually independent conditional on θ.

(iii) $f_{W_1|W_2}(w_1|w_2)$ and $f_{\theta|W_1}(\theta|w_1)$ form a bounded complete family of distributions indexed by W_2 and W_1, respectively.

▸ Definition

(iv) For all $\bar{\theta}, \tilde{\theta} \in \Theta$, the set $\{w_3 : f_{W_3|\theta}(w_3|\bar{\theta}) \neq f_{W_3|\theta}(w_3|\tilde{\theta})\}$ has positive probability whenever $\bar{\theta} \neq \tilde{\theta}$.

(v) We normalize $\mu_1(\theta) = \theta$ and $E[\varepsilon_t|\theta] = 0$ for all t.

Lemma 1: Identification in the Simple Case

Lemma 1

Under Assumption 1, $f_\theta(\cdot)$, $f_{\varepsilon_t}(\cdot)$, and $\mu_t(\cdot)$ are identified $\forall t$.

Proof:

- Thm 1 of Hu and Schennach (2008) gives identification of $f_{W_1|\theta}(\cdot|\cdot)$, $f_{W_2|\theta}(\cdot|\cdot)$, and $f_{W_3,\theta}(\cdot,\cdot)$ from $f_{W_1,W_2,W_3}(\cdot,\cdot,\cdot)$
- $f_\theta(\cdot)$ can be recovered from $f_{W_3,\theta}(\cdot,\cdot)$ by integrating out W_3 (Cunha, Heckman and Schennach 2010)
- identify $\mu_2(\cdot)$ and $\mu_3(\cdot)$ from $E[W_t|\theta] = \mu_t(\theta)$ given $f_{W_t|\theta}(\cdot|\cdot)$
- $f_{\varepsilon_t}(\cdot)$ is identified from $f_{\varepsilon_t}(\varepsilon) = f_{W_t|\theta}(\mu_t(\theta) + \varepsilon)$, since $\mu_t(\cdot)$ and $f_{W_t|\theta}(\cdot|\cdot)$ are already known.

Nonparametric Identification: Serially Correlated Shocks

Now, consider heteroskedastic permanent shocks and an $MA(1)$ process for ε_{it}:

$$\begin{aligned} W_{it} &= \mu_t(\theta_i) + \kappa_{i,t} + \nu_{i,t} \\ \kappa_{i,t} &= \kappa_{i,t-1} + \eta_{i,t} \\ \nu_{i,t} &= \xi_{i,t} + \beta_t \xi_{i,t-1} \end{aligned}$$

- Allow $\eta_{i,t} = \sigma_t(\theta_i)\zeta_{i,t}$
- Identification for most parameters/densities/functions requires $T \geq 9$ (we focus on $T = 9$)
- Use differences to eliminate correlations in shocks

Generalizing Lemma 1

Taking first differences and looking at observations far enough apart, we can get back to independence and apply Lemma 1 (or something similar):

$$
\begin{aligned}
W_1 &= \theta + \varepsilon_1 = \theta + \{\eta_1 + \nu_1\} \\
\Delta W_4 &= \Delta\mu_4(\theta) + \Delta\varepsilon_4 = \Delta\mu_4(\theta) + \{\eta_4 + \Delta\nu_4\} \\
\Delta W_7 &= \Delta\mu_7(\theta) + \Delta\varepsilon_7 = \Delta\mu_7(\theta) + \{\eta_7 + \Delta\nu_7\}
\end{aligned}
$$

- repeat for other triplets $(W_2, \Delta W_5, \Delta W_8)$ and $(W_3, \Delta W_6, \Delta W_9)$ to identify all $\mu_t(\cdot)$
- identification now comes from relationship between earnings and future earnings changes

Assumption 2

The following conditions hold for $T = 9$:

(i) The joint density of θ, W_1, W_2, W_3, $\Delta W_4, \ldots, \Delta W_9$ is bounded and continuous, and so are all their marginal and conditional densities. $f_\theta(\cdot)$ is non-vanishing on $\mathbb{R}$.

(ii) Unobserved components ζ_t, ξ_t, and θ are mutually independent for all $t = 1, \ldots, 9$.

(iii) $f_{W_t|\Delta W_{t+3}}(w_t|\Delta w_{t+3})$ and $f_{\theta|W_t}(\theta|w_t)$ form a bounded complete family of distributions indexed by ΔW_{t+3} and W_t, respectively, for $t = 1, 2, 3$.

(iv) For all $\bar{\theta}, \tilde{\theta} \in \Theta$ and $t = 7, 8, 9$, the set $\{w : f_{\Delta W_t|\theta}(w|\bar{\theta}) \neq f_{\Delta W_t|\theta}(w|\tilde{\theta})\}$ has positive probability whenever $\bar{\theta} \neq \tilde{\theta}$.

(v) We impose the following normalizations: $\kappa_0 = \xi_0 = 0$, $\mu_1(\theta) = \theta$, $E[\zeta_t] = E[\xi_t] = 0$, $E[\zeta_t^2] = 1$, and $\sigma_t(\cdot) > 0$ for all t.

(vi) For all t, we assume the Carleman's condition holds for ζ_t and ξ_t.

Theorem 1: Identification with Serially Correlated Errors

Theorem 1

Under Assumption 2, $f_\theta(\cdot), \{f_{\eta_t}(\cdot), f_{\xi_t}(\cdot), \beta_t\}_{t=1}^{7}$, *and* $\{\mu_t(\cdot)\}_{t=1}^{9}$ *are identified.*

Proof has three steps:

- Identify $f_\theta(\cdot)$ and $\mu_t(\cdot)$ for all t using Lemma 1. Details
- Identify $E[\xi_t^2]$, β_t, and $\sigma_t(\cdot)$ for $t = 1, ..., 7$ using various second moments.
- Identify $f_{\eta_t}(\cdot)$ and $f_{\xi_t}(\cdot)$ for $t = 1, ..., 7$.

Some General Comments on Identification

- For $T \geq 9$, this general strategy can be used to identify $f_\theta(\cdot)$, $\{\mu_t(\cdot)\}_{t=1}^{T}$ and $\{f_{\eta_t}(\cdot), f_{\xi_t}(\cdot), \beta_t\}_{t=1}^{T-2}$
- If shocks are all homoskedastic, cannot have flat regions in $\Delta\mu_t(\cdot)$ for $t = 7, 8, 9$
- Identification approach rules out an autoregressive process where transitory shocks never die out
- Can handle arbitrarily long $MA(q)$ process, but may require a long panel
  - $MA(q)$ requires $T \geq 6 + 3q$ time periods
  - can only identify shock process through $T - q - 1$

Moment-Based Approach

Now, consider a moment-based approach

- Assume $\mu_t(\theta) = m_{0,t} + m_{1,t}\theta + ... + m_{p,t}\theta^p$
- Normalize $\mu_1(\theta) = \theta$ and $E[\mu_t(\theta)] = 0$ for $t = 2, ..., T$
- Assume $f_{\xi_t}(\cdot)$ and $f_{\eta_t}(\cdot)$ are time-specific
- Assume θ, η_t and ξ_t are mutually independent with η_t and ξ_t independent over time
- Normalize $E[\theta] = E[\eta_t] = E[\xi_t] = 0$

Using Variances & Covariances

Assuming shocks begin at age $a = 1$, we have variances:

$$E[W^2_{i,a,t}|a,t] = \sum_{j=0}^{p}\sum_{j'=0}^{p} m_{j,t}m_{j',t}E[\theta^{j+j'}] + \sum_{j=0}^{a-1}\sigma^2_{\eta_{t-j}} + \sigma^2_{\xi_t} + \sum_{j=1}^{\min\{q,a-1\}} \beta^2_{j,t}\sigma^2_{\xi_{t-j}}$$

and covariances (for $l \geq 1$):

$$E[W_{i,a,t}W_{i,a+l,t+l}|a,t,l] = \sum_{j=0}^{p}\sum_{j'=0}^{p} m_{j,t}m_{j',t+l}E[\theta^{j+j'}] + \sum_{j=0}^{a-1}\sigma^2_{\eta_{t-j}} + E(\nu_{i,a,t}\nu_{i,a+l,t+l})$$

Number of Moments and Parameters for $MA(1)$

Consider the number of moments & parameters for one cohort using variances & covariances

Number of parameters:

- $2p - 1$ parameters for $E[\theta^2]$,...,$E[\theta^{2p-1}]$,$E[\theta^{2p}]$
- $(p + 1)(T - 1)$ parameters for $\mu_t(\theta)$ polynomials, $t = 2, ..., T$
- $2(T - 2)$ parameters for $\sigma^2_{\eta_t}$ and $\sigma^2_{\xi_t}$, $t = 1, ..., T - 2$
- $T - 3$ parameters for β_t, $t = 2, ..., T - 2$
- Total number of parameters: $(4 + p)T + p - 9$

Number of moments:

- $\frac{T(T+1)}{2}$ variance/covariance terms
- $T - 1$ moments coming from $E[\mu_t(\theta)] = 0$, $t = 2, ..., T$
- Total moments: $\frac{T(T+1)}{2} + T - 1$

Identification

Necessary condition for identification: $T \geq 3$ and $p \leq \frac{T^2 - 5T + 16}{2(T+1)}$

- Cubic $\mu_t(\cdot)$ requires $T \geq 10$
- Adding higher residual moments can be helpful
- Higher moments necessary to identify higher moments of shock distributions $f_{\eta_t}(\cdot)$ and $f_{\xi_t}(\cdot)$

Multiple Cohorts

- With changing distribution of cohorts over time (aging in and out of panel), it is important to account for the fact that older cohorts have accumulated a longer history of shocks
  - we assume shocks start at age 20
- Additional cohorts can aid in identification, since $\mu_t(\cdot)$ does not vary across cohorts

PSID Data: Overview

- PSID is a longitudinal survey of a representative sample of US individuals and their families
- Collected annually through 1997, biennially starting in 1999
- We use data from interview years 1971 through 2009
- Earnings are collected for the previous year, so data cover calendar years 1970-2008
- Earnings: household head's total wages and salaries (excluding farm and business income)
- Earnings reported in 1996 dollars using CPI-U-RS

Sample Restrictions

- Core (SRC) sample with nonzero weights
  - exclude oversamples (SEO, Latino) and nonsample persons
- Male heads of households
- Ages 30-59
- Positive annual wages and weeks worked
- Non-students
- Trim top and bottom 1% of wages within each age-year cells (ten-year age group used)
- Resulting data set has 3,302 men and 33,207 person-year observations

Sample Statistics

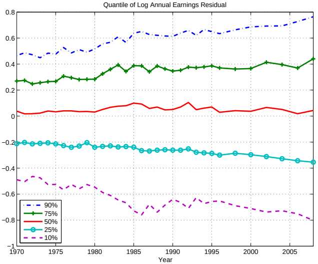
- Race: 92% White, 6% black, 1% hispanic
- Age: mean age is 47
- Educational Attainment

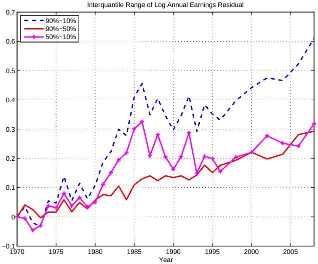
Education (years)	Percent
Elementary (1-5)	1.2
Middle (6-8)	5.0
Some High (9-11)	9.9
Completed High (12)	33.7
Some College (13-15)	20.0
Completed College (16)	20.6
Advanced Degrees (17+)	9.8

Obtaining Residuals

- We focus on the distribution of residual earnings, controlling for differences in education, race, and age
- Run a cross-sectional regression of log earnings for each year on
  - age dummies
  - race dummies
  - education dummies
  - race dummies $\times$ cubic polynomial in age
  - education dummies $\times$ cubic polynomial in age

Residual Earnings Inequality in the US, 1970-2008





Moment-Based Estimation

- Assume $\beta_{j,t} = \beta_j$ for all j and t
- We assume $\sigma^2_{\eta_\tau} = \sigma^2_{\eta_0}$ and $\sigma^2_{\xi_\tau} = \sigma^2_{\xi_0}$ for all τ years prior to our data
  - other assumptions yield similar results
- Assume homoskedastic shocks for most of the analysis
- Use minimum distance for estimation
  - aggregate into three age categories for variance/covariance moments
  - weight moments by share of observations used for that moment

Decomposition

We focus on decomposing variance of earnings residuals into three components:

- pricing of unobserved skills: $Var[\mu_t(\theta)]$
- permanent shocks: $\sigma^2_{\kappa_t} = \sum_{j=0}^{a-1} \sigma^2_{\eta_{t-j}}$
- transitory shocks: $\sigma^2_{\nu_t} = \sigma^2_{\xi_t} + \sum_{j=1}^{q} \beta^2_{jt} \sigma^2_{\xi_{t-j}}$

Estimation with linear $\mu_t(\theta)$

We begin by assuming $\mu_t(\theta)$ is linear

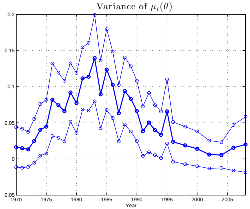
- only use variances & covariances in estimation
- begin by assuming distribution of θ is the same across cohorts, but explore changes across cohorts later

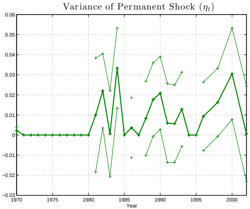
MD estimation

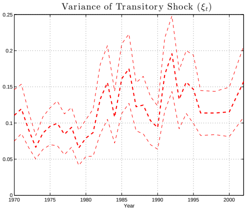
	Constant μ_t	Time-Varying $\mu_t(\cdot)$			
	$MA(3)$	$MA(1)$	$MA(2)$	$MA(3)$	$MA(5)$
Min. Obj. Fun.	**168.27**	130.73	124.16	**121.10**	116.93
β_1	0.326	0.297	0.281	0.288	0.299
	(0.027)	(0.033)	(0.026)	(0.027)	(0.027)
β_2	0.222	·	0.186	0.172	0.194
	(0.025)	·	(0.025)	(0.021)	(0.020)
β_3	0.246	·	·	0.141	0.137
	(0.034)	·	·	(0.025)	(0.022)
β_4	·	·	·	·	0.126
	·	·	·	·	(0.020)
β_5	·	·	·	·	0.084
	·	·	·	·	(0.024)

Variance of θ and shocks

MA(3) model with time-varying $\mu_t(\theta)$

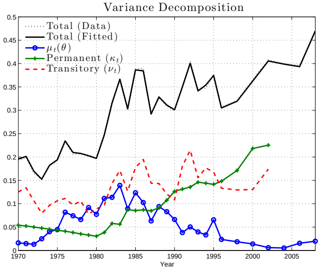


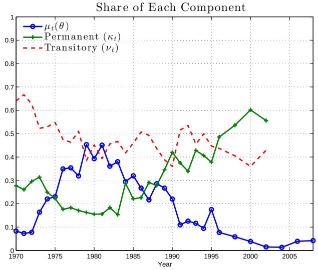




Variance Decomposition

MA(3) model with time-varying $\mu_t(\theta)$

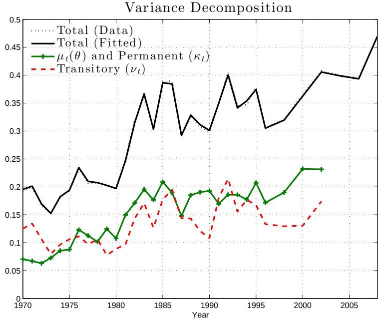


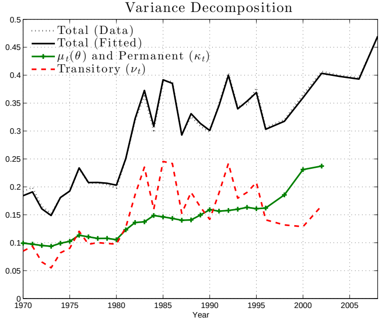


Comparison with time-invariant μ model

MA(3) shocks with time-varying vs. time-invariant $\mu_t(\theta)$

Time-Varying $\mu_t(\theta)$	Time-Invariant $\mu(\theta)$

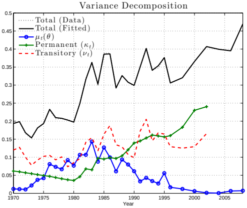




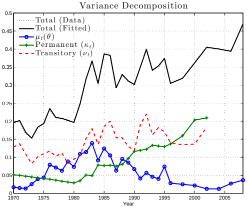
Other Transitory Processes

MA(1), MA(5), and ARMA(1,1) shocks with time-varying $\mu_t(\theta)$

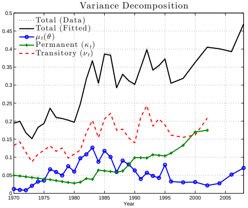
MA(1)



MA(5)



ARMA(1,1)



η_t as Skill Shocks

Suppose we interpret η_t as skill shocks so

$$W_{it} = \mu_t(\theta_i + \kappa_{it}) + \nu_{it}$$

where we assume linear $\mu_t(\cdot)$

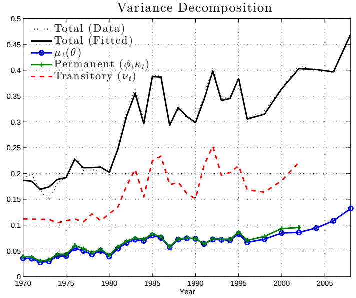
η_t as Skill Shocks

Suppose we interpret η_t as skill shocks so

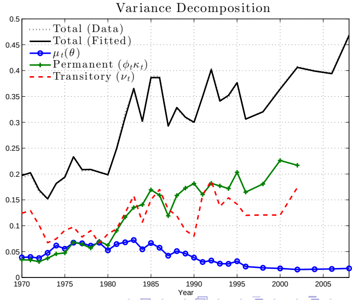
$$W_{it} = \mu_t(\theta_i + \kappa_{it}) + \nu_{it}$$

where we assume linear $\mu_t(\cdot)$

Constant $\sigma^2_{\eta t}$



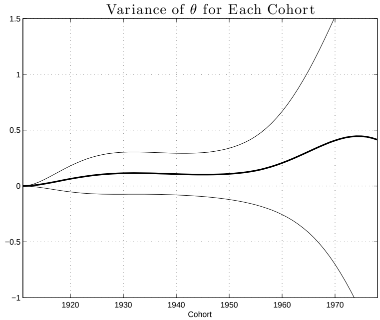
Time-Varying $\sigma^2_{\eta t}$

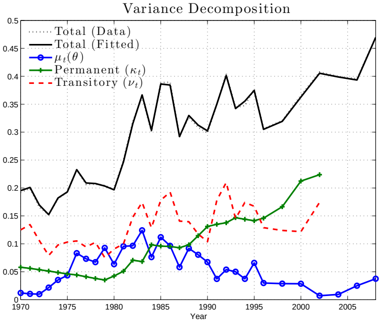


Allowing for cohort differences in distribution of θ

- Shifts in mean of θ are absorbed in age and time effects before obtaining residuals
- We examine whether the variance of θ varies across cohorts
  - birth cohorts from 1911 to 1978
  - assume a cubic spline in year of birth with two interior knots

Cohort Differences in Variance of θ





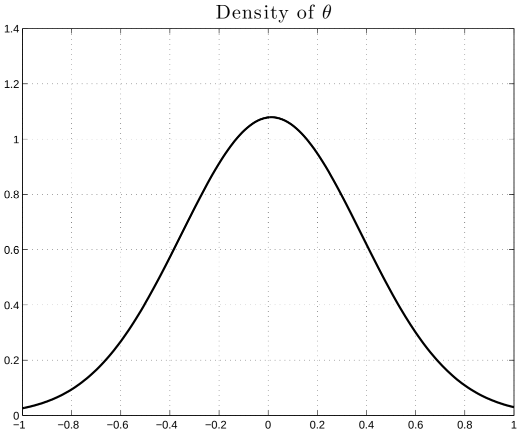
Estimation with cubic $\mu_t(\theta)$

Now, assume

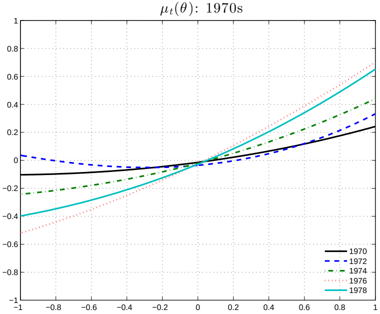
- $\mu_t(\theta)$ are time-varying cubic functions with $\mu_{1985}(\theta) = \theta$
- $f_\theta(\cdot)$ is a mixture of two normals (same for all cohorts)
- permanent and MA(3) transitory shocks

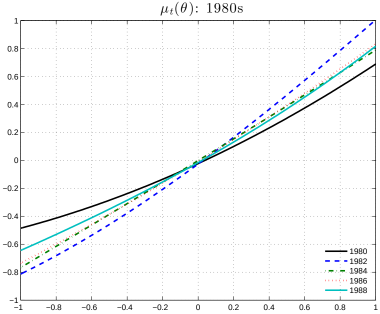
We now use all second- and third-order moments of log earnings residuals

Distribution of θ (1985)

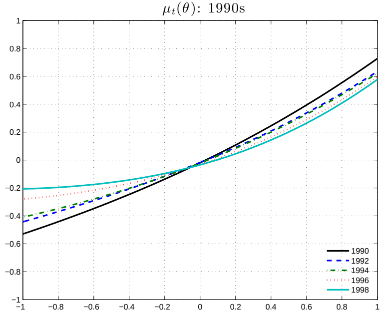


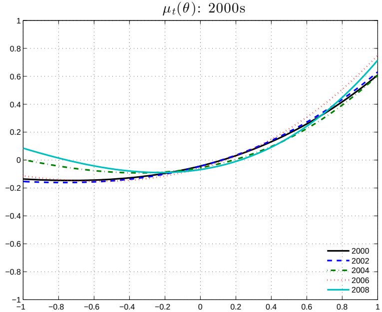
Estimated $\mu_t(\theta)$ functions





Estimated $\mu_t(\theta)$ functions

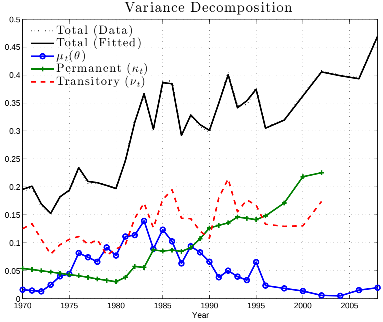




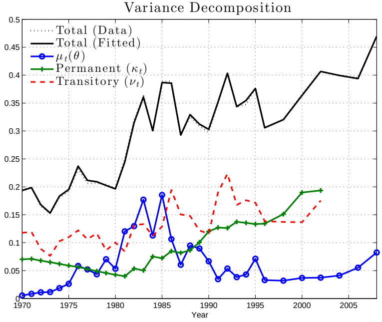
Variance Decomposition

Comparison with linear $\mu_t(\theta)$

linear $\mu_t(\theta)$

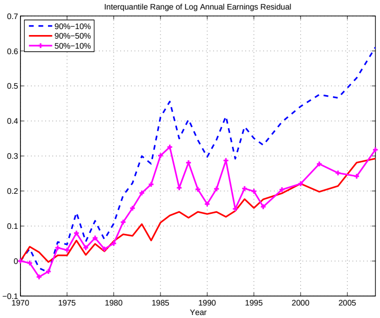


cubic $\mu_t(\theta)$

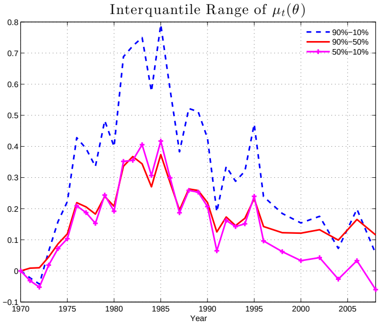


Evolution of $\mu_t(\theta)$ distribution vs. residual distribution

Residuals

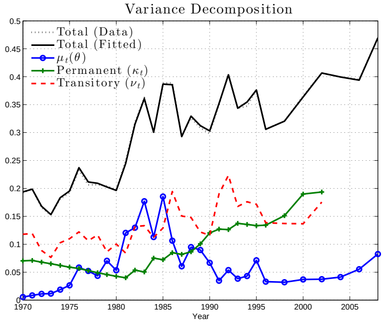


$\mu_t(\theta)$

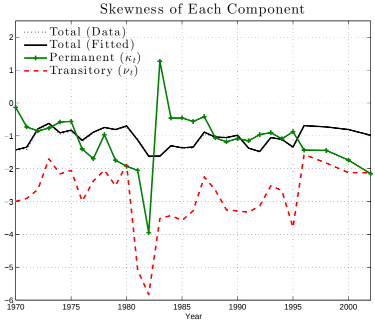


Variances and Skewness Over Time

Variance

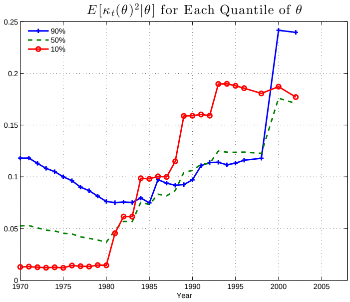


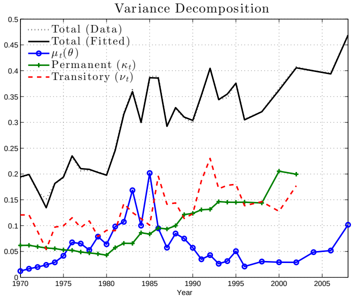
Skewness



Heteroskedasticity in Permanent Shocks

Consider $\sigma_t(\theta) = s_{0,t} + s_{1,t}\theta$.





Summary & Conclusions

- We consider identification and estimation of a model with unobserved skill differences and time-varying
  - skill 'pricing' functions
  - permanent shocks
  - transitory shocks
- Identification
  - prove nonparametric identification
  - discuss identification for a moment-based approach
- Estimation
  - Minimum Distance estimation using second- and third-order residual moments
  - Use male log earnings residuals in PSID

Summary & Conclusions

- Results suggest that all components of earnings have played an important role since 1970
  - ‘returns’ to unobserved skill increased broadly in 1970s and early 1980s but fell in late 1980s/early 1990s
    - stronger decline in value of unobserved skill at bottom than top after 1995 (partial polarization)
  - variance of unobserved skills changed little over 1925-55 cohorts
  - variance of transitory shocks jumped in early 1980s and bounced around after
  - variance of permanent shocks rose consistently over 1980s and 1990s (especially among low ability)
- Inequality in unobserved skills evolves quite differently from overall residual inequality
  - accounting for idiosyncratic shocks is important for understanding role of unobserved skills

An Economic Explanation

A theory based on slow diffusion of skill-biased technology in frictional labor markets can be helpful (based on Violante 2002)

- introduction of skill-biased technology in 1970s
- quick diffusion to most high-skill workers
  - increase in $\mu_t'(\cdot)$ but not $\sigma^2_{\eta_t}$
- by mid-1980s, skill begins to diffuse more slowly to lower skilled workers
  - decrease in $\mu_t'(\cdot)$
- matching of low-skill workers to new technologies random due to market frictions
  - increase in $\sigma^2_{\eta_t}(\theta)$ for low ability workers
- rising $\sigma^2_{\eta_t}$ consistent with rising variance of wages paid across firms (e.g., Dunne et al. 2004, Barth et al. 2011)

Future Efforts

- Estimate differences by education and/or race
  - allow $f_\theta(\cdot)$ and $\mu_t(\theta)$ to vary by education/race
  - what roles do changes in unobserved skill distributions and pricing play in earnings gaps?
- Move from modelling residuals to earnings/wages themselves
- Allow for multiple unobserved skills
- Examine implications for consumption inequality within and across cohorts over time

Thanks

Bounded complete family of distributions

$f_{\theta|W}(\theta|W)$ forms a bounded complete family of distributions indexed by W if $g(\theta) = 0$ is the only bounded function that solves:

$$\int g(\theta) f_{\theta|W}(\theta|W) d\theta = 0, \quad \forall W$$

- Standard assumption in nonparametric identification literature related to invertability of conditional expectation integral function
- E.g. violated if
  - $\theta \perp\!\!\!\perp W$, since $g(\theta) = \theta - E(\theta)$ solves the equation above
  - $f_{\theta|W}$ is symmetric about 0 for all W, e.g. W only affects variance of θ

Back

Step 1: Details on identification of μ_2, μ_3,...

Consider the second subset of equations:

$$\begin{aligned}
W_2 &= \mu_2(\theta) + \varepsilon_2 = \theta_2 + \{\eta_1 + \eta_2 + \nu_2\} \\
\Delta W_5 &= \Delta\mu_5(\theta) + \Delta\varepsilon_5 = g_5(\theta_2) + \{\eta_5 + \Delta\nu_5\} \\
\Delta W_8 &= \Delta\mu_8(\theta) + \Delta\varepsilon_8 = g_8(\theta_2) + \{\eta_8 + \Delta\nu_8\}
\end{aligned}$$

where $g_t(\theta_2)$ is implicitly defined by $\Delta\mu_t(\theta) = g_t(\mu_2(\theta))$.

- Can identify $f_{\theta_2}(\cdot)$, $g_5(\cdot)$, and $g_8(\cdot)$ using same approach
- Recover the function $\mu_2(\cdot)$ by $\mu_2(\theta) = F_{\theta_2}^{-1}(F_\theta(\theta))$
- Once we identify $\mu_2(\cdot)$, $\Delta\mu_5(\cdot)$ and $\Delta\mu_8(\cdot)$ are identified from $\Delta\mu_t(\theta) = g_t(\mu_2(\theta))$

Back