Global Income Dynamics Database Project: Brazil

Niklas Engbom (NYU)
Gustavo Gonzaga (PUC-Rio)
Christian Moser (Columbia & FRB Minneapolis)
Roberta Olivieri (Cornell)

November 20, 2020

Introduction

But First...

...a big ‘Thank You’ to the organizers,

and to Serdar Ozkan & Sergio Salgado, who

did a fantastic job with coding and support!

Income dynamics in a large developing country: Brazil

Part I: Formal sector (administrative data)

- Large initial inequality, decline over time, except very top
- Bottom-driven decline in earnings volatility
- Significant increase in upward mobility, particularly at top

Part II: Formal & informal sectors (household surveys)

- More volatile income in informal sector
- Large decline in volatility of income in Brazil since 2000
- Accounted for by within-sector decline & formalization

Data

1. Relação Anual de Informações Sociais (RAIS)
   - Linked employer-employee panel data
   - Covers all formal workers 1994–2017, large sample during high-inflation period 1985–93
   - Covers app. 1.2 billion job records, or 40 million per year
   - Key variables: worker ID, gender, age, education, job start and end month, monthly earnings

2. Pesquisa Mensal de Emprego (PME)
   - Household survey rotating panel data (4-8-4 months)
   - Covers formal, informal, and nonemployed workers in six metropolitan regions 2002–2015
   - Covers app. 7.3 million interviews, or 500 thousand per year
   - Key variables: worker ID, gender, age, education, monthly earnings, employment status, work permit

Sample Selection

- Ages 25–55
- Separate analysis by gender and pooled
- Earnings include wage, salary, and bonus payments
  - Censored earnings above 120 times MW in RAIS, $\approx 0.01\%$
  - Uncensored earnings in PME household survey data
- Compute total annual earnings, y_{it}, for person i in year t
- Drop if $y_{it} \leq \frac{1}{2} \times 40 \frac{\text{hours}}{\text{week}} \times 4 \frac{\text{weeks}}{\text{month}} \times 3\text{months} \times MW_t$

Summary Statistics on Final Sample

Table: Formal sector (RAIS) data – Heterogeneity Sample

	Real annual earnings (2018 USD)					
Year	Mean	Std. dev.	P10	P50	P90	Observations
2002	30,956	42,042	6,552	17,335	66,913	14,200,000
2012	28,469	36,260	8,313	16,959	58,608	22,500,000

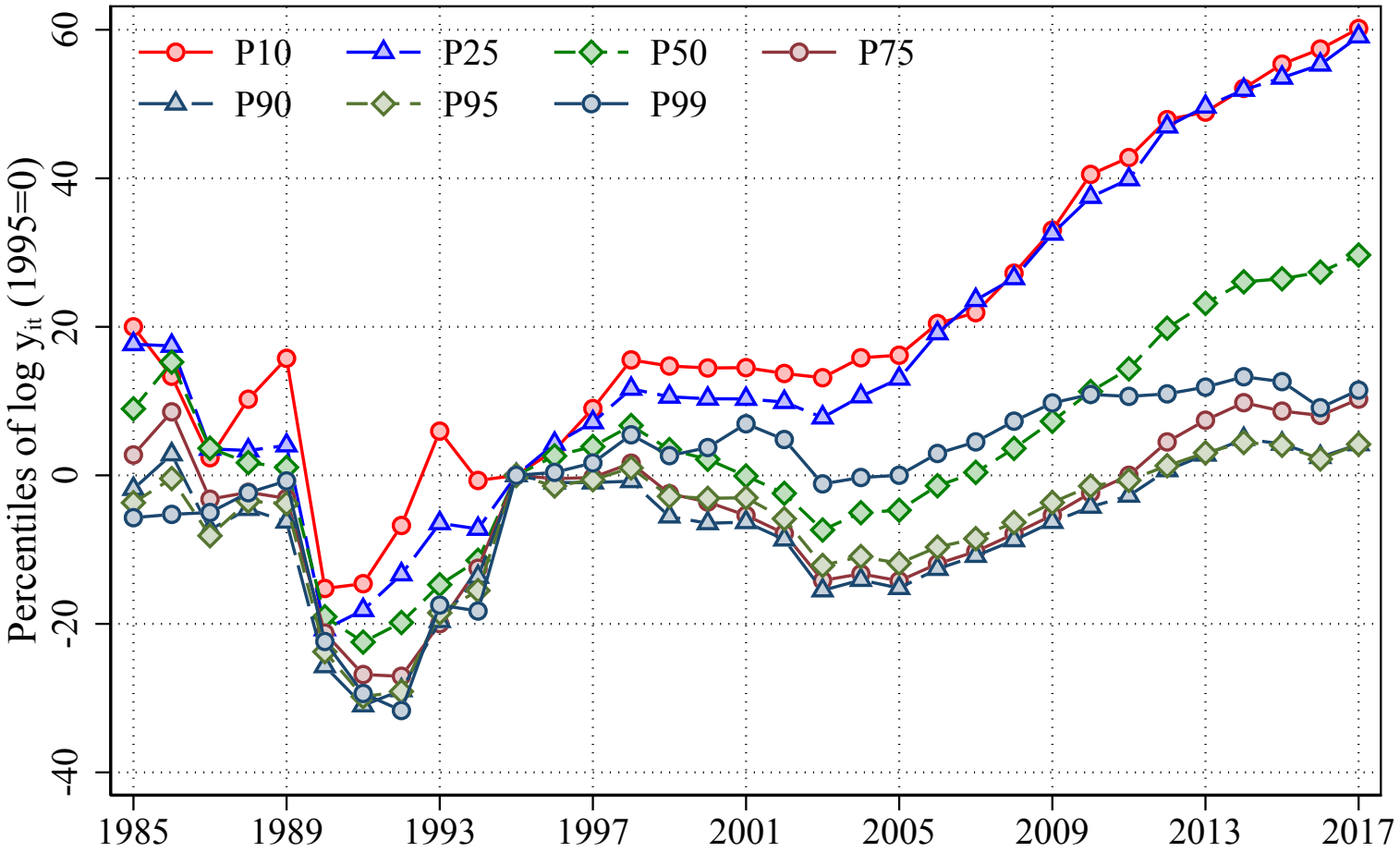
Part I: Formal Sector

Road Map for Part I

1. Inequality and Concentration
2. Volatility and Higher-Order Moments
3. Mobility

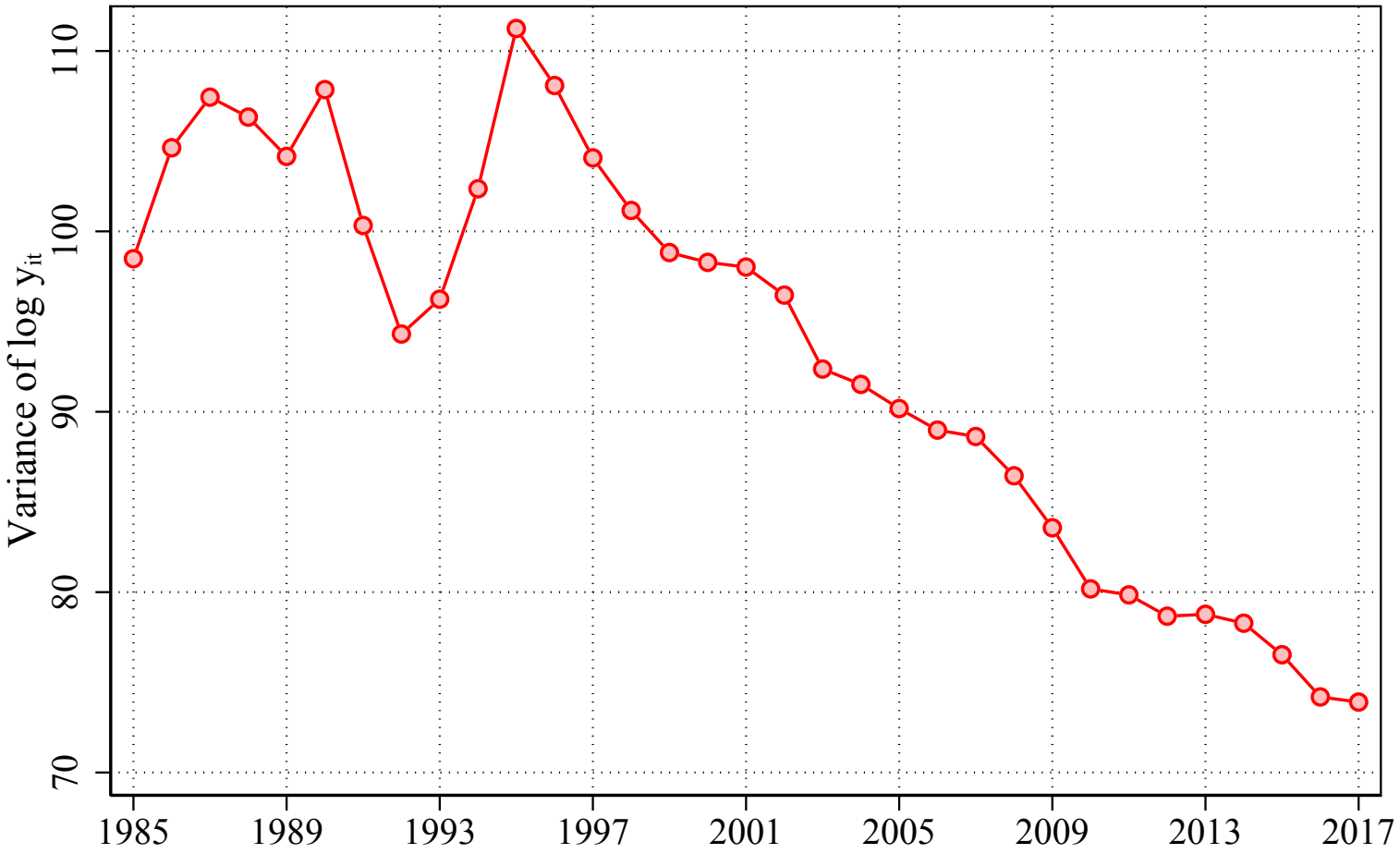
Normalized Log Earnings Percentiles ▸ Permanent ▸ Residual

- After 1995, fast earnings growth, particularly at bottom



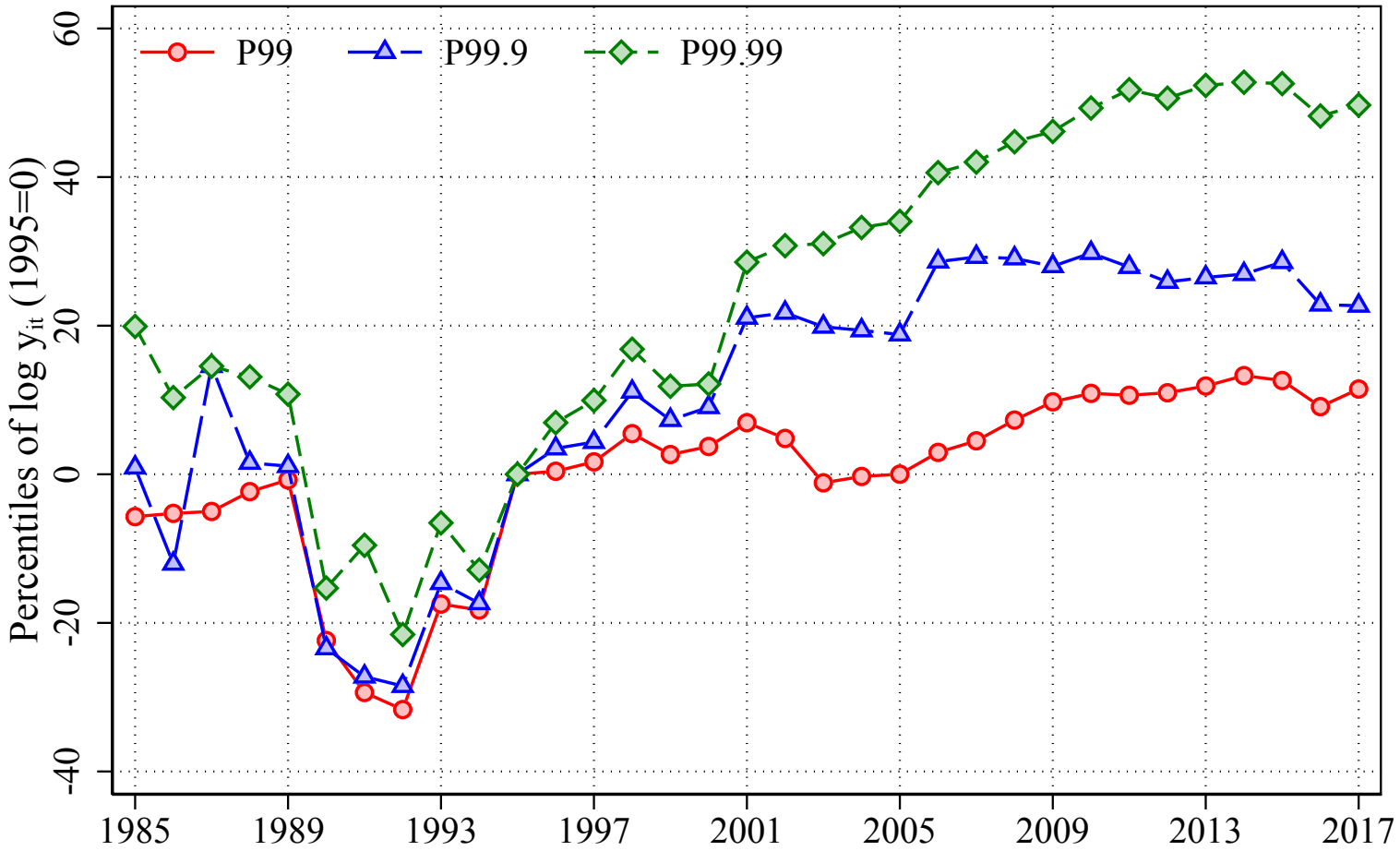
Variance of Log Earnings ▸ Permanent ▸ Residual

- Initially high but decreasing inequality (Alvarez et al. '18)



Top Log Earnings Percentiles

- Large inequality decline, but top earnings grew fast >P99



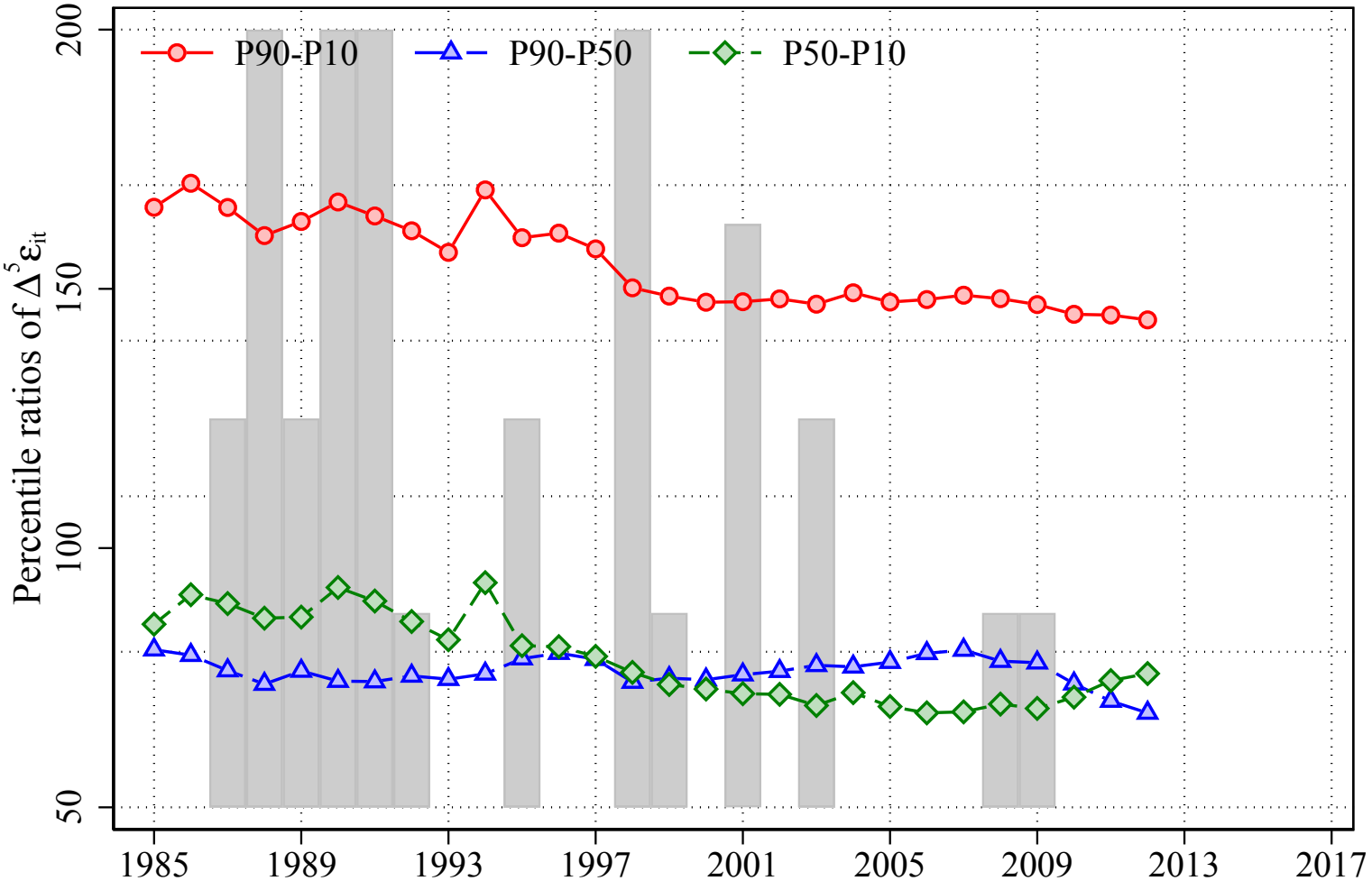
Road Map for Part I

1. Inequality and Concentration ✓
2. Volatility and Higher-Order Moments
3. Mobility

Log Perc. Ratios of 5-Year Resid. Earnings Change

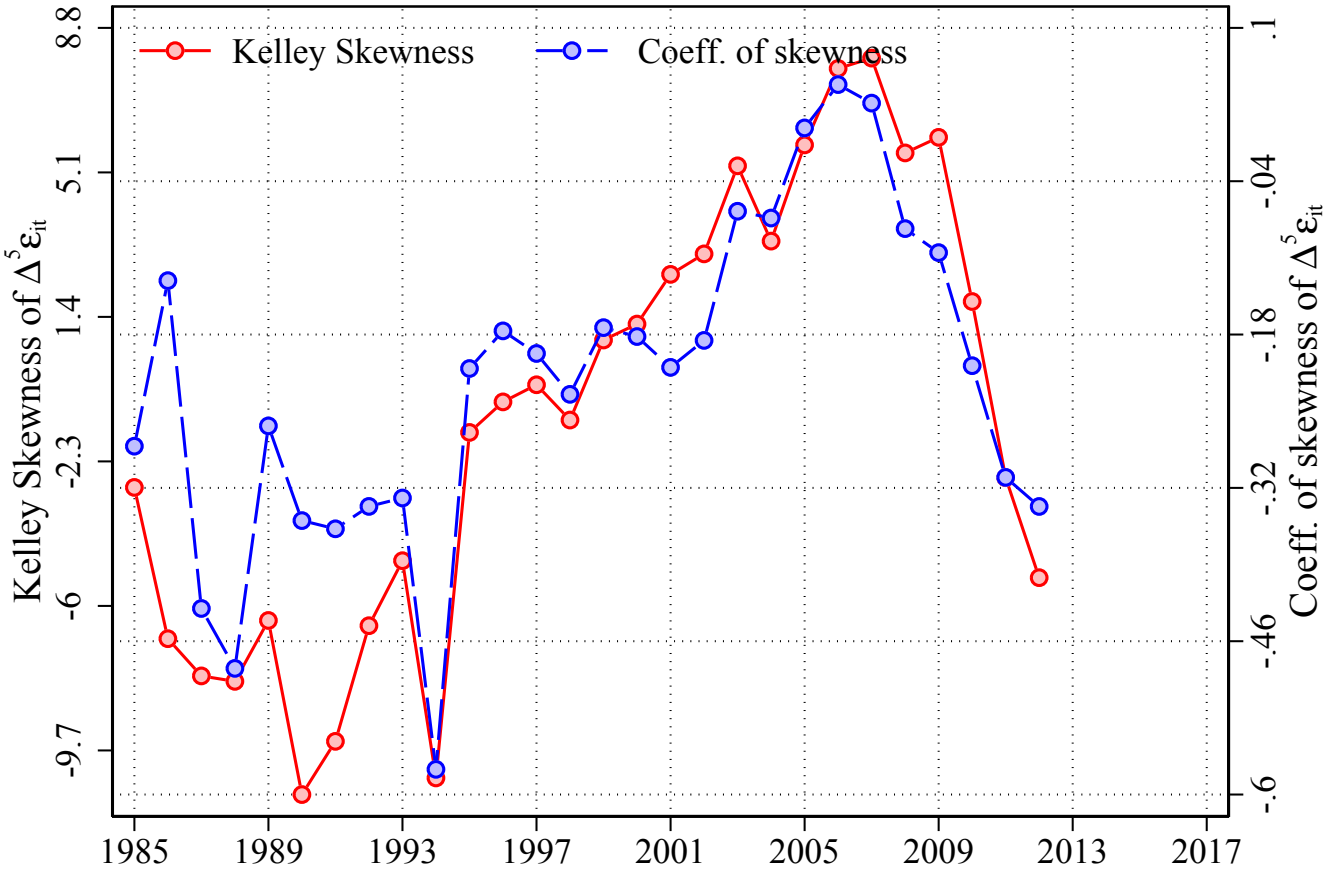
▸ Short-Term

- Decline in earnings volatility
  - Compression in upper tail from 1985–1990 and 2007–2012
  - Compression in lower tail from 1994–2000



Skewness of 5-Year Residualized Earnings Change

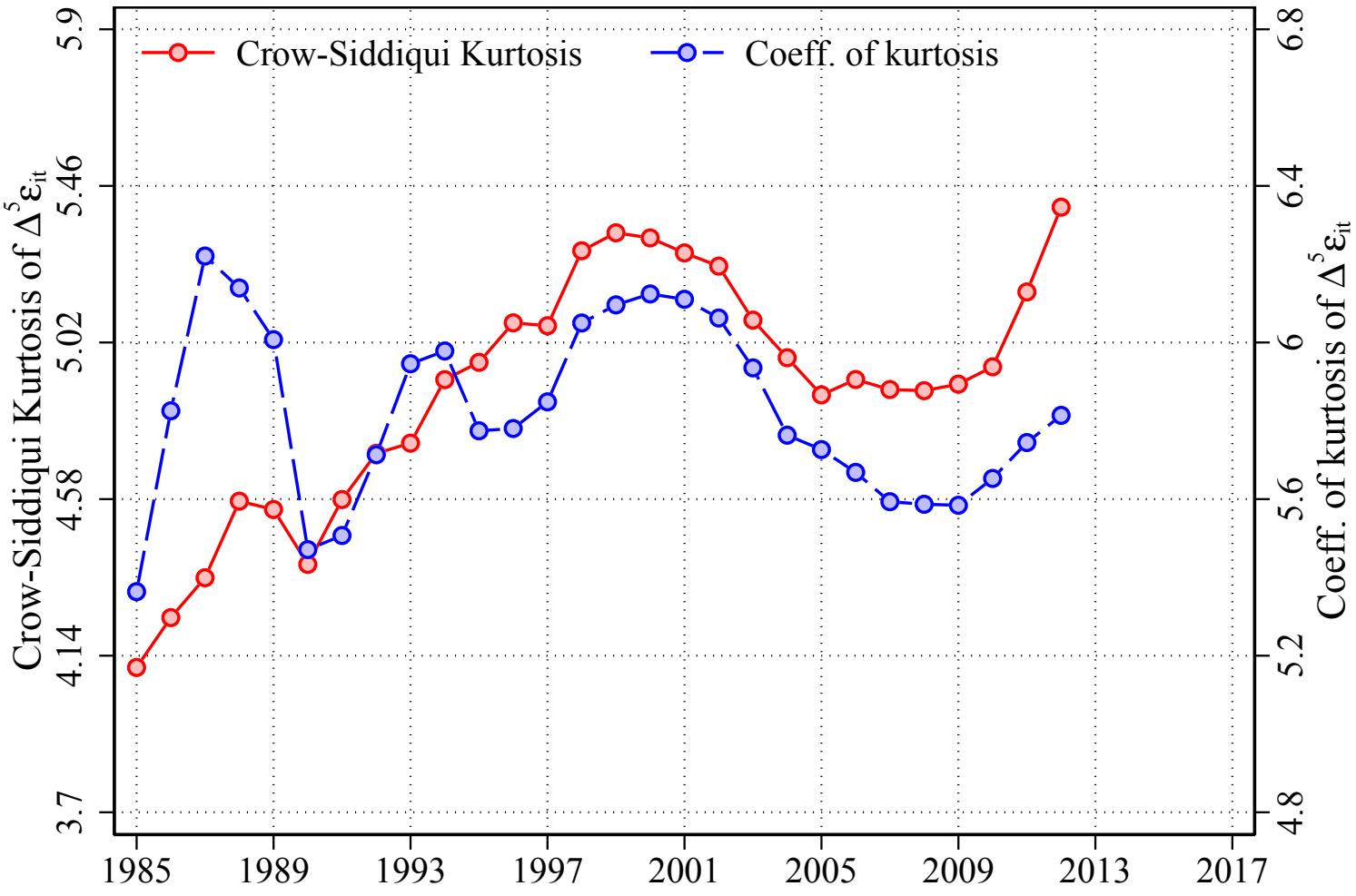
- Slightly higher skewness rel. to US (Guvenen et al., '19)
- Some evidence of procyclicality of skewness
  - 1994–2009: rise during economic expansion
  - 1994 and 2009–2011: dip in years before big recessions



Kurtosis of 5-Year Residualized Earnings Change

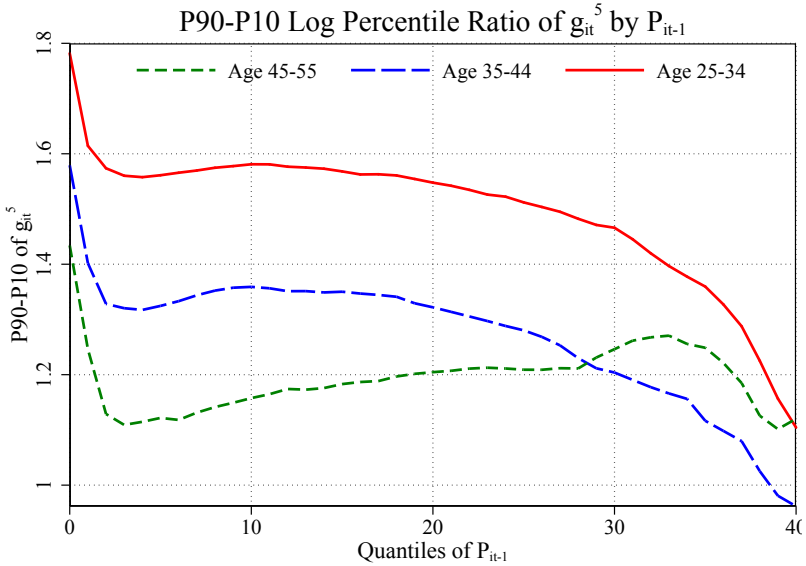
▸ Short-Term

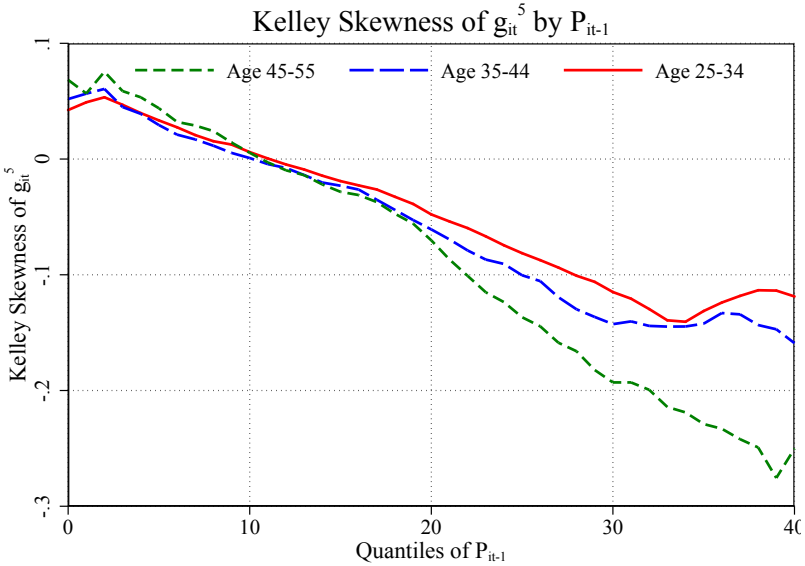
- Slightly lower kurtosis rel. to US (Guvenen et al., '19)
- No clear cyclical pattern, but maybe increasing over time

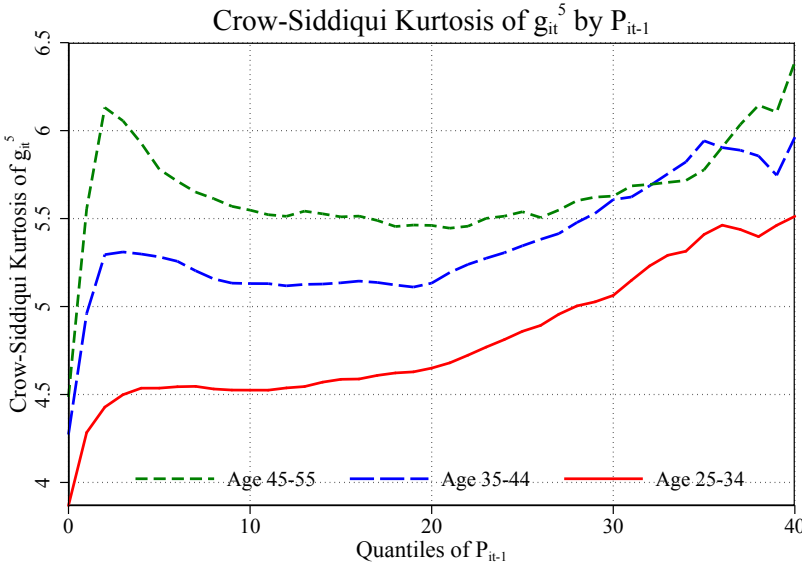


5-Year Resid. Earnings Changes, by Age & Perm. Inc.

- 2nd/3rd/4th moment decr./decr./incr. in perm. income







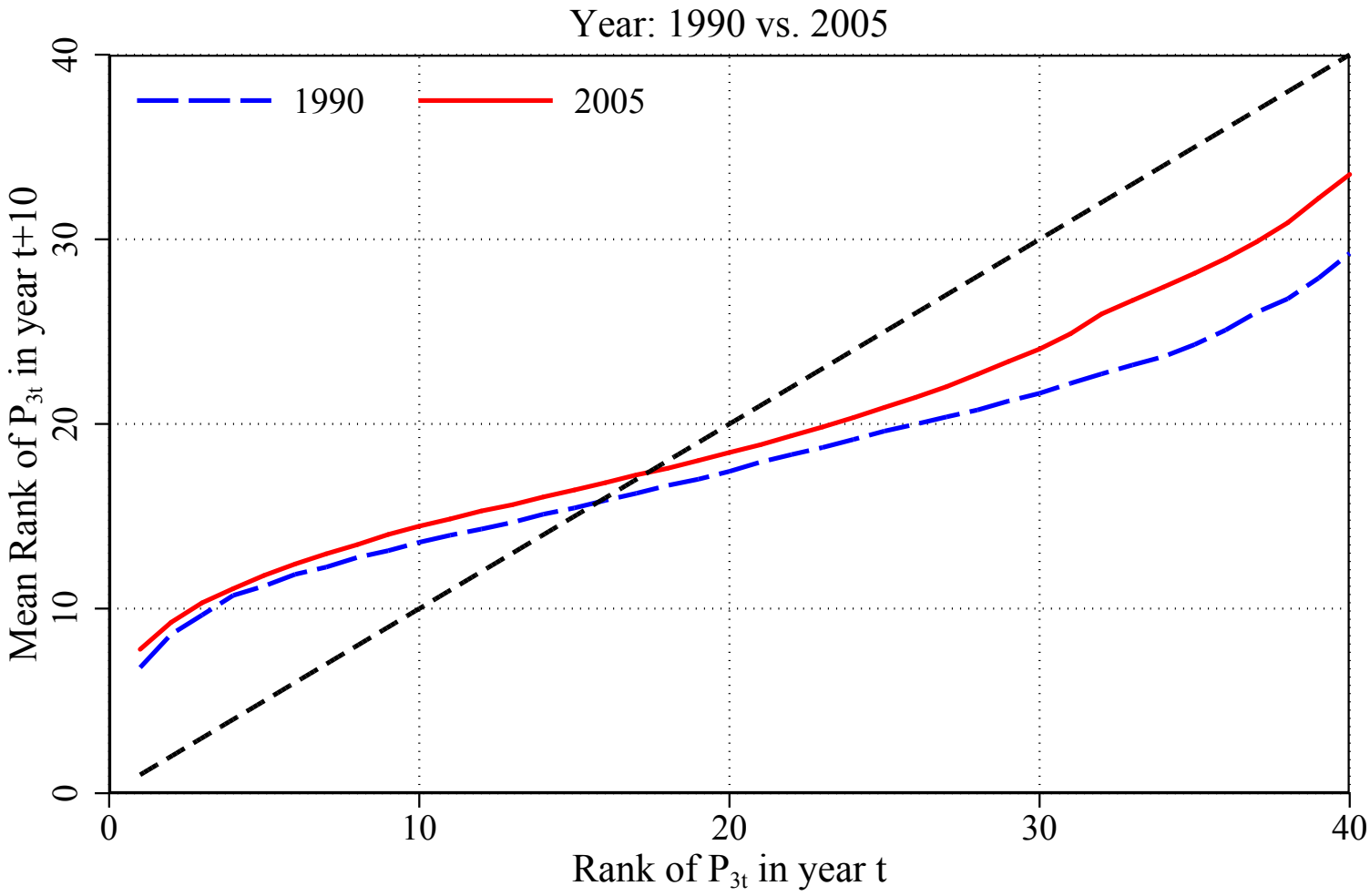
Road Map for Part I

1. Inequality and Concentration ✓
2. Volatility and Higher-Order Moments ✓
3. Mobility

Permanent Mobility in Residual Income, 1990 vs. 2005

▸ More

- Significant increase in upward mobility, particularly at top



Note: Higher mean future ranks at all current ranks is possible due to switchers into and out of the population.

Part II: Formal & Informal Sectors

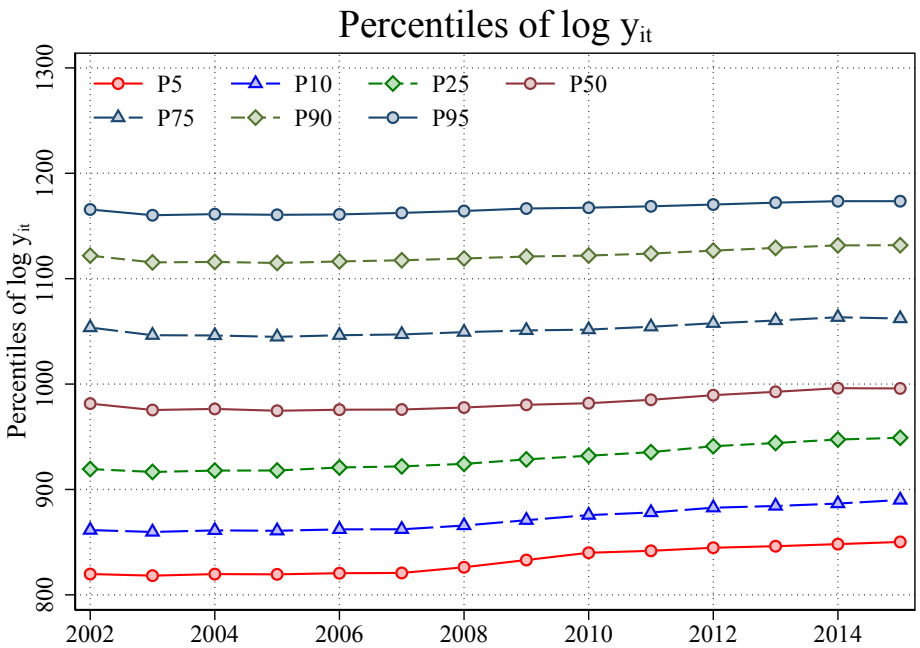
Road Map for Part II

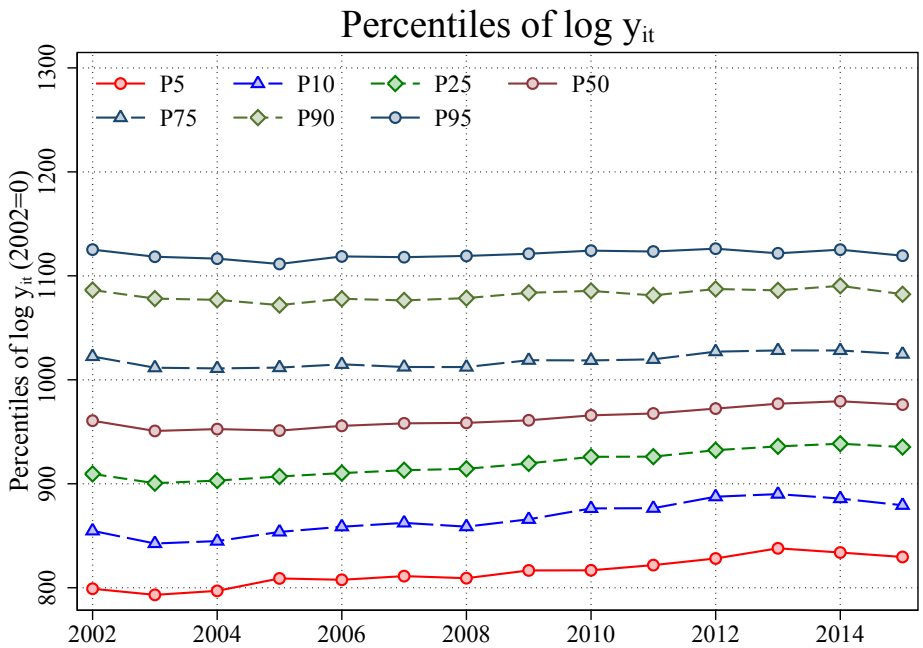
1. Validation: Comparing Admin. & Survey Data
2. Informal Dynamics & Sector Transitions: 5 Facts

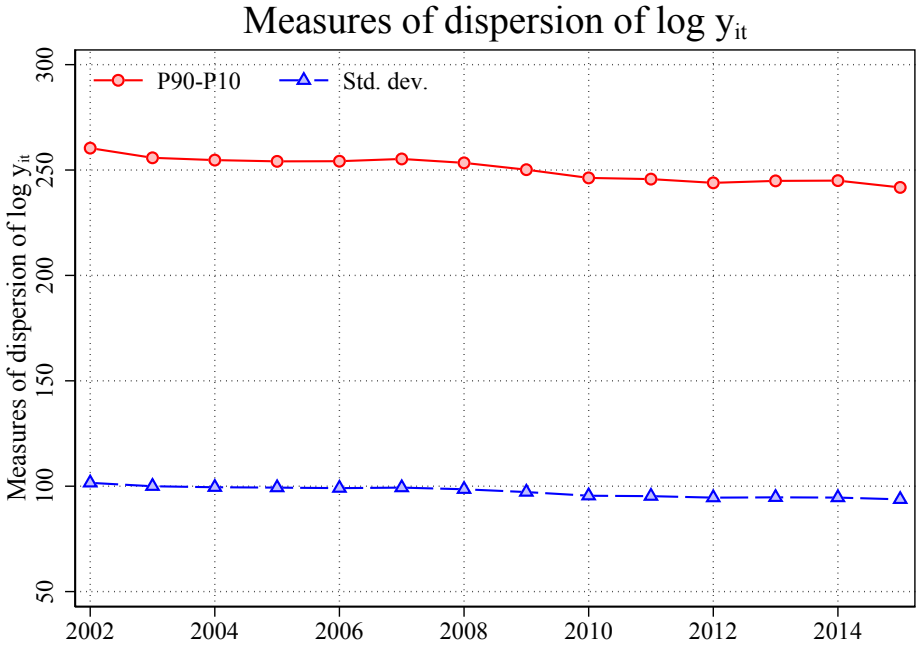
Validation Exercise

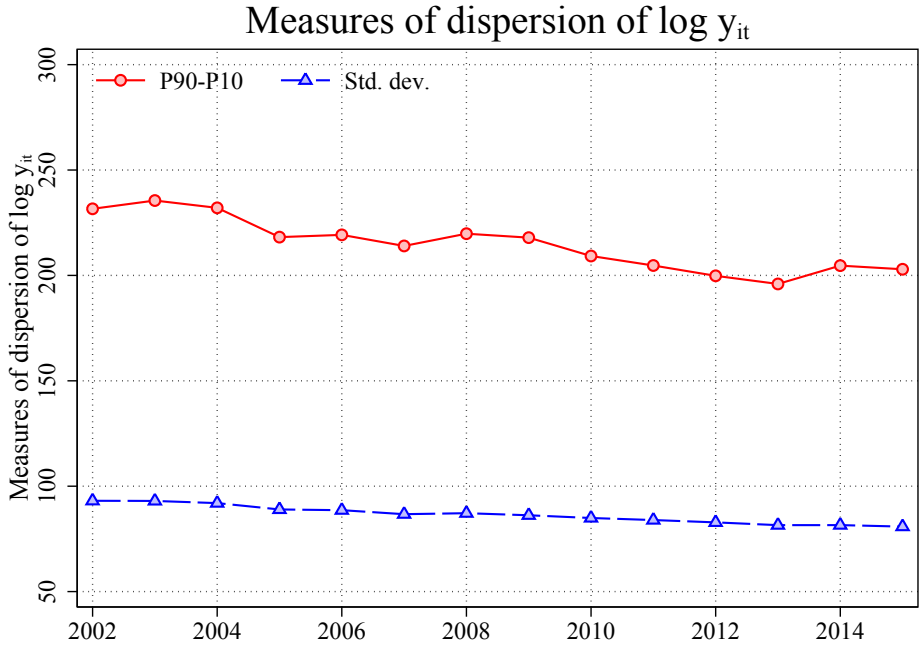
- How comparable are formal-sector earnings (dynamics) in admin. data vs. household survey data?
- Make two datasets as comparable as possible:
  - Restrict admin. data to 6 largest metro regions covered in household survey data
  - Apply same upper-income censoring in household survey data as in admin. data

Inequality in Admin. (LHS) & Survey (RHS) Data ▸ Normalized

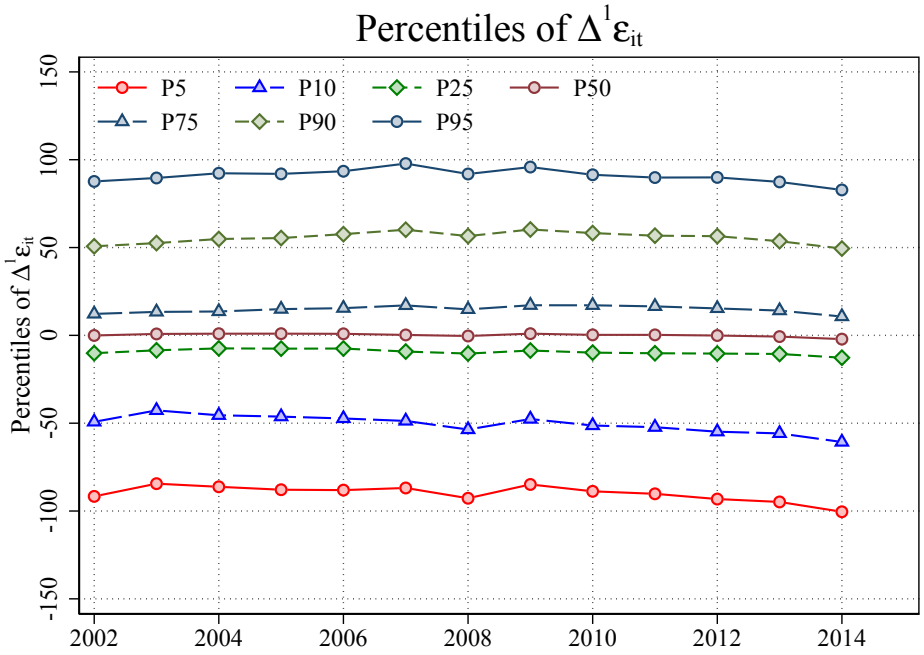


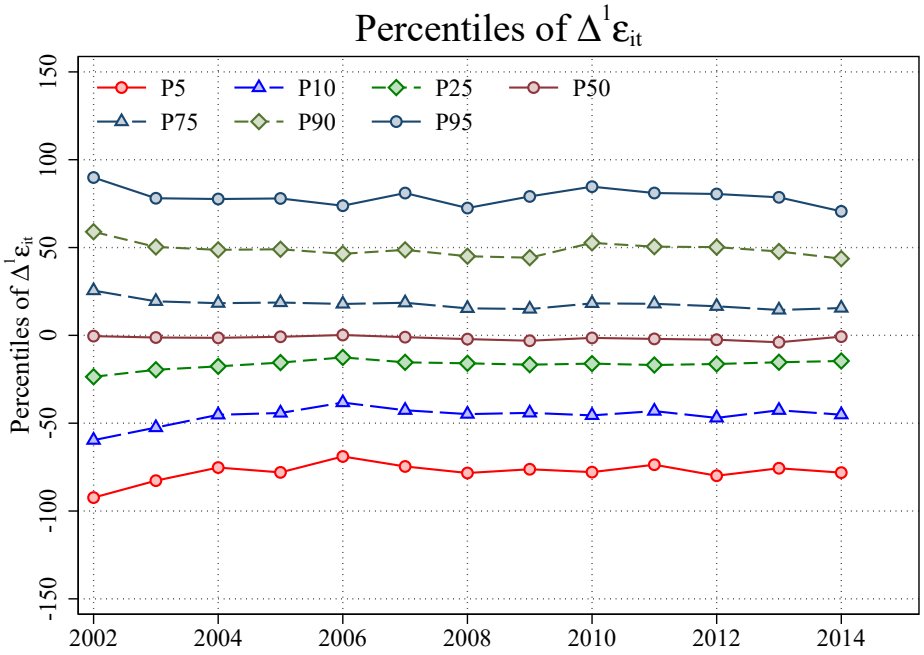


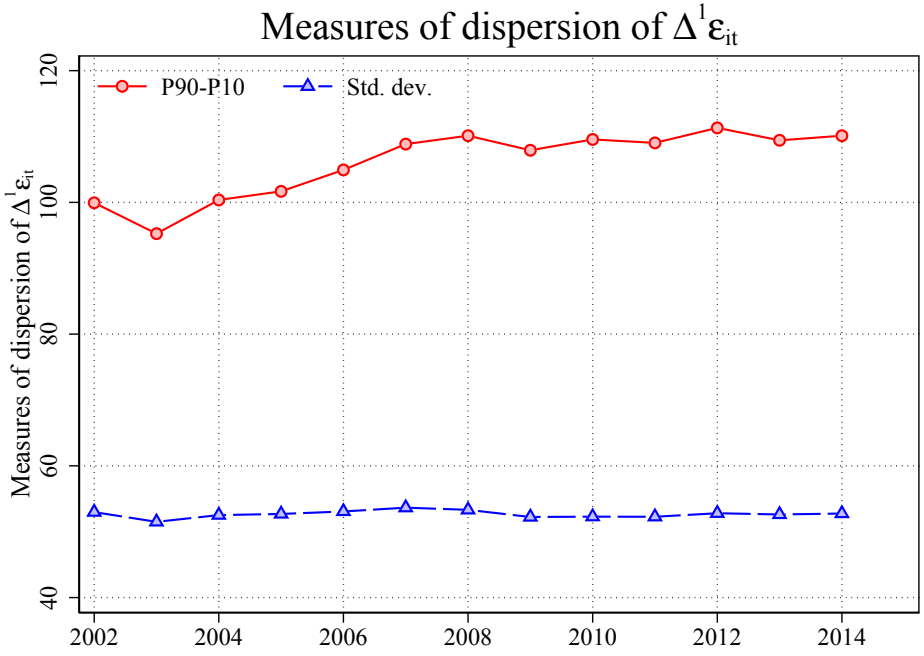


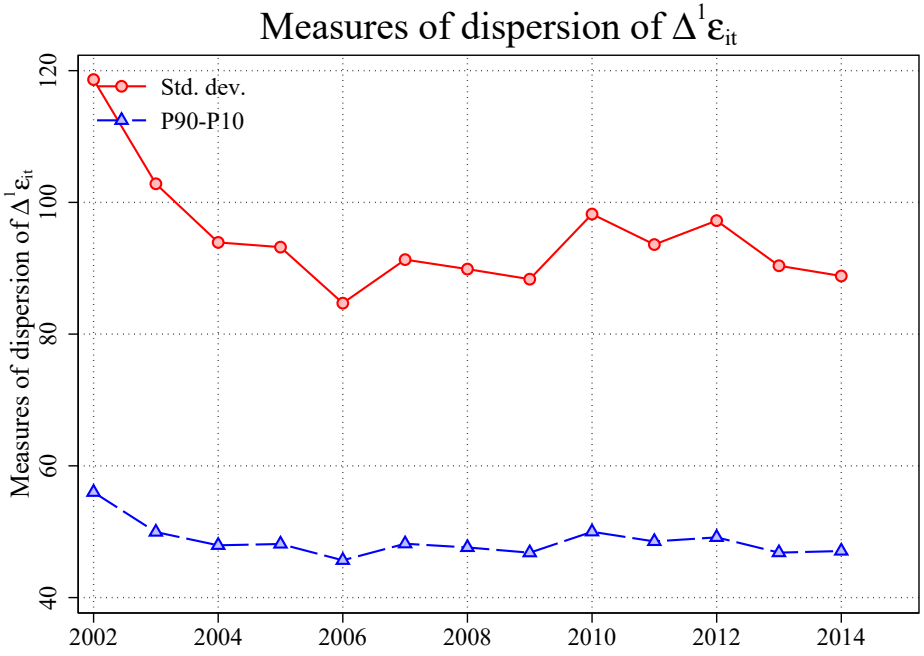


Volatility in Admin. (LHS) & Survey (RHS) Data ▸ Normalized







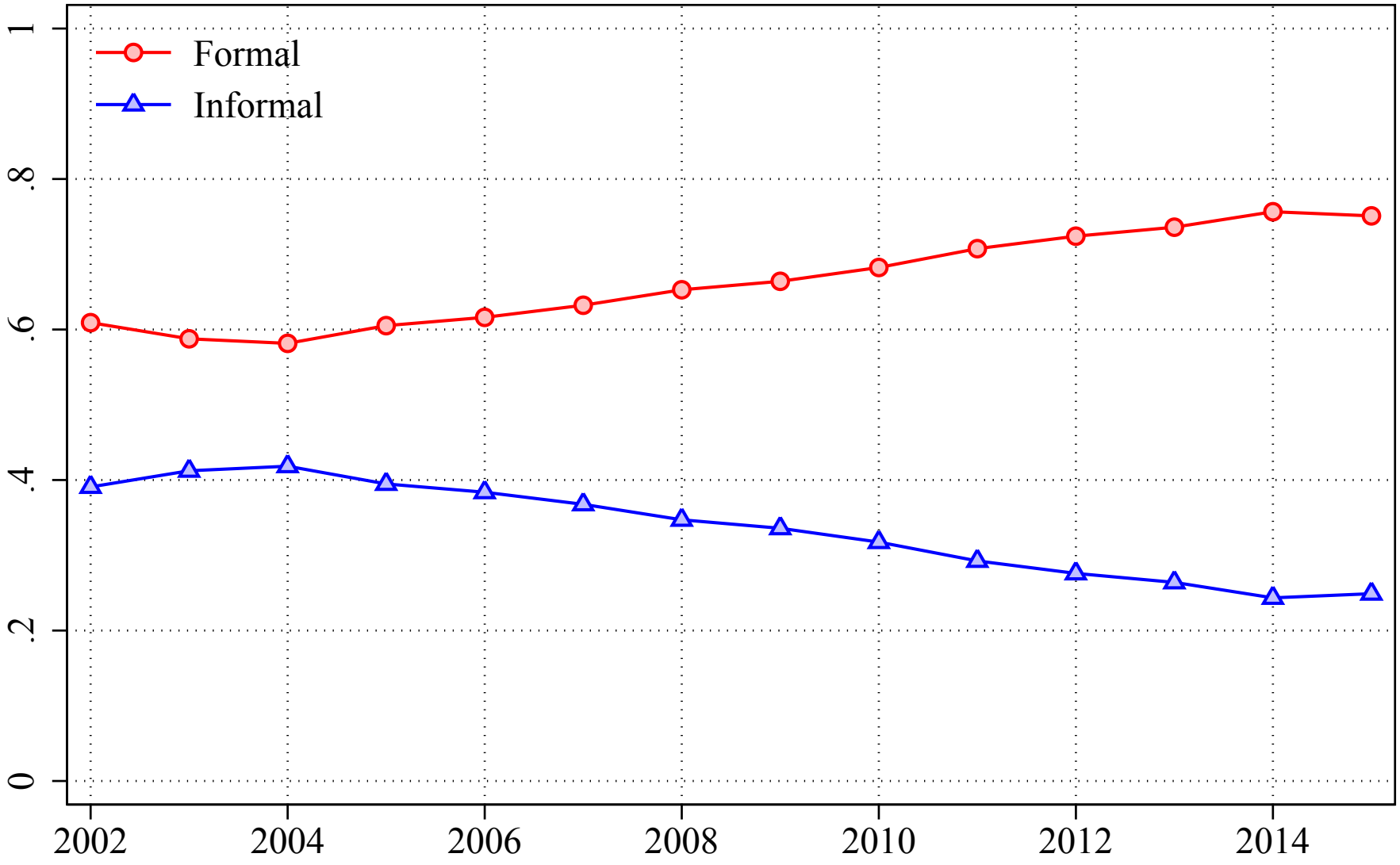


Road Map for Part II

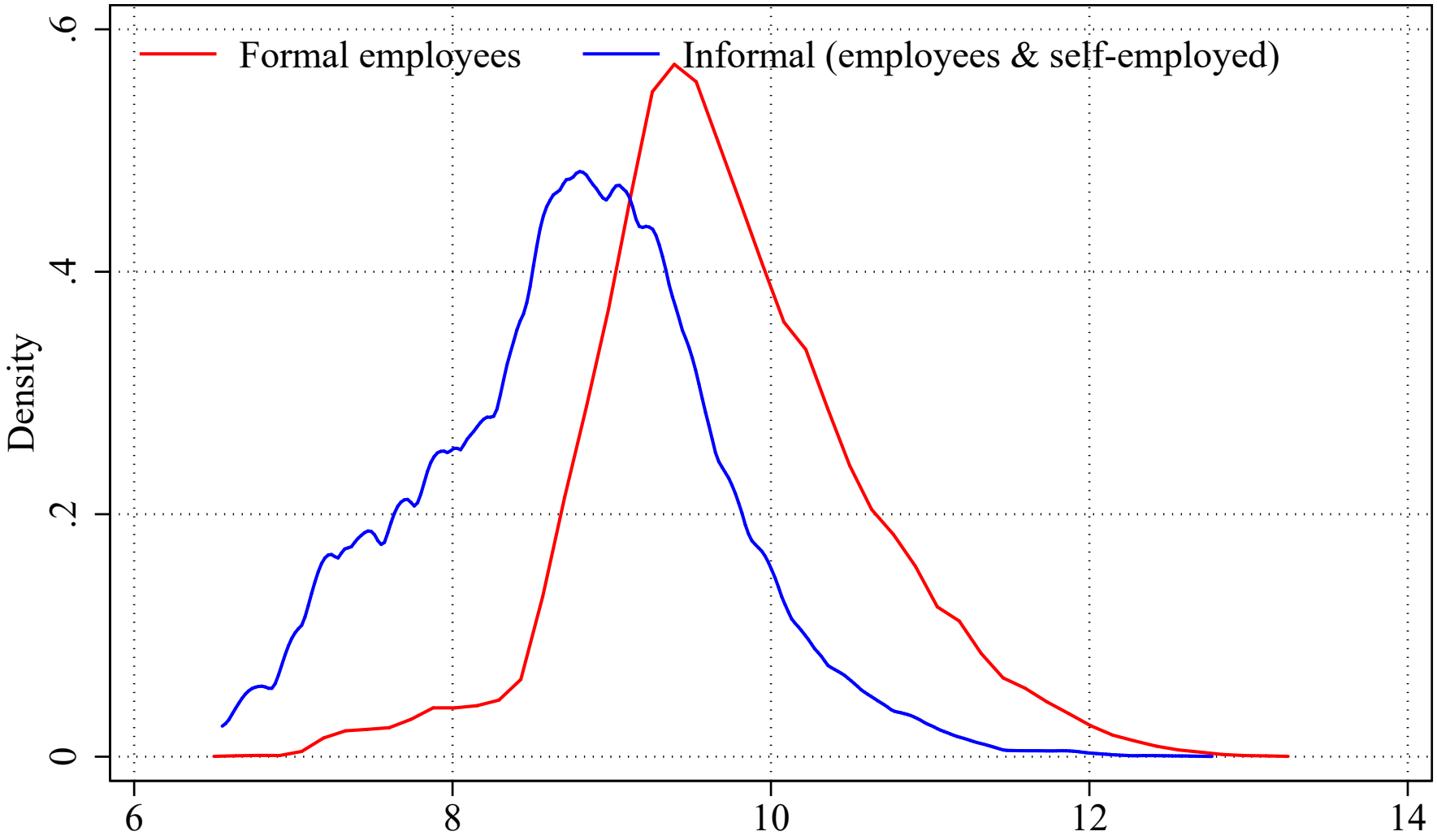
1. Validation: Comparing Admin. & Survey Data ✓

2. Informal Dynamics & Sector Transitions: 5 Facts

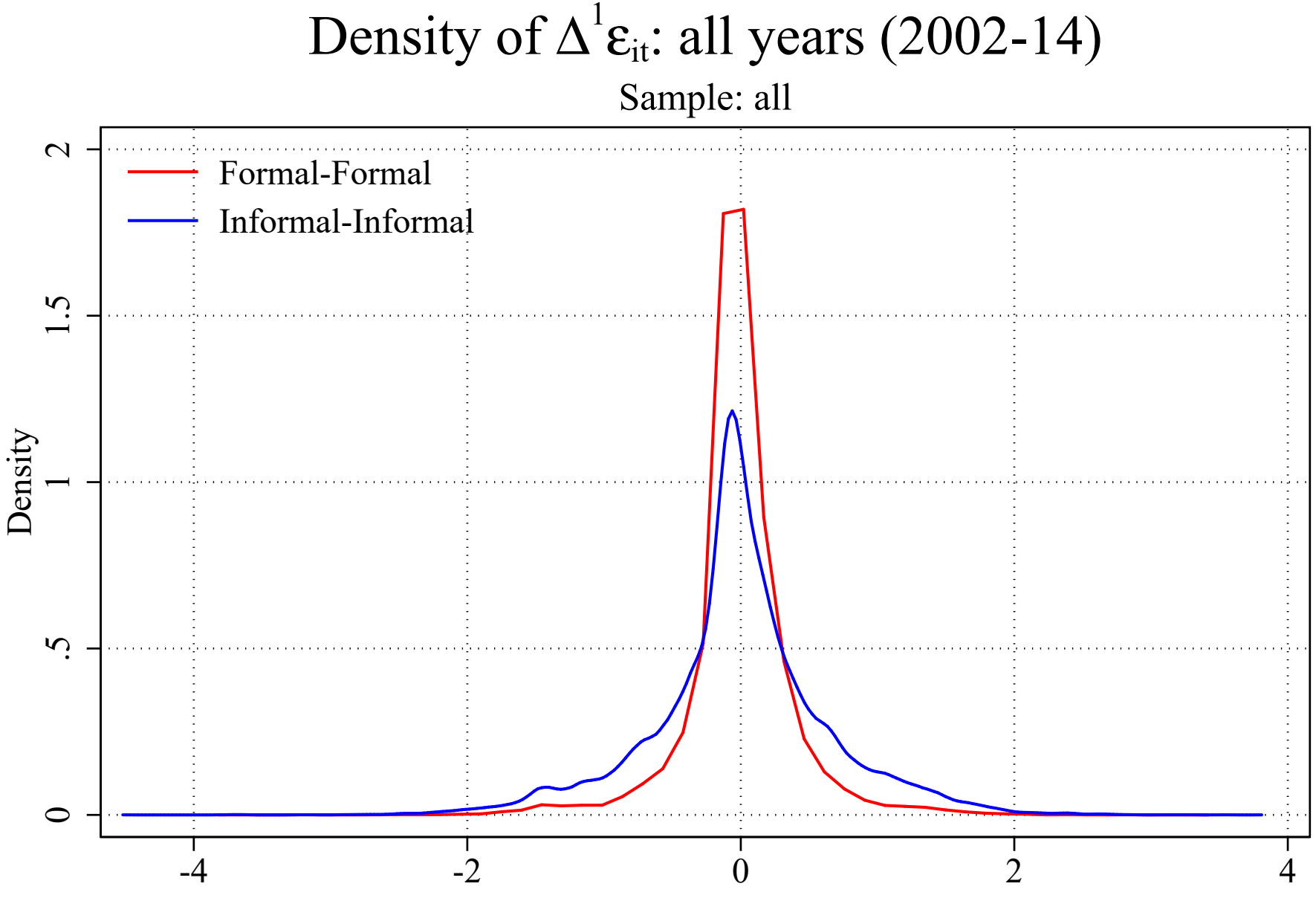
Formal vs. Informal Sector: Employment Shares



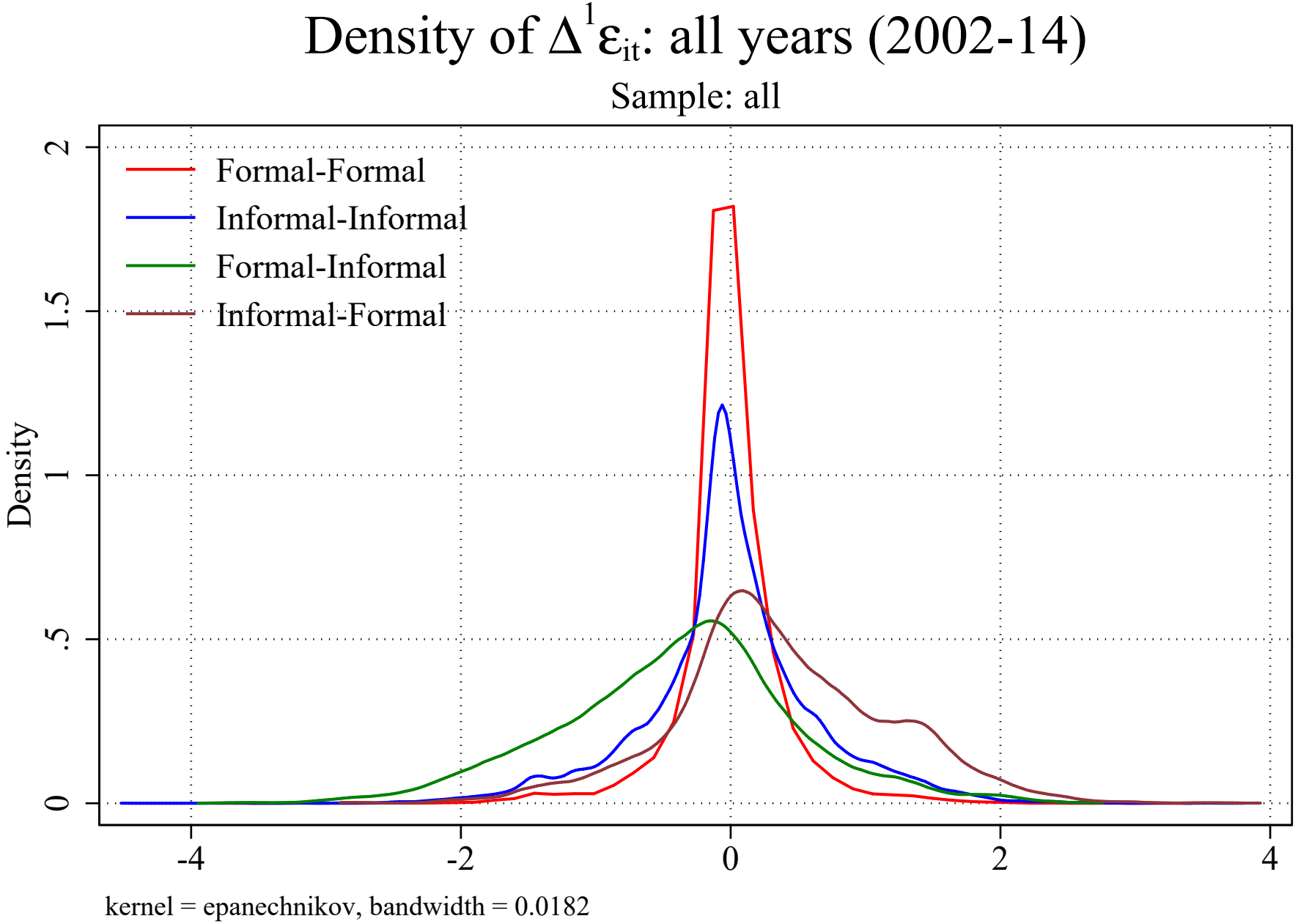
Formal vs. Informal Sector Earnings in 2003



Fact I: Same Average, Higher Volatility in Informal Sector



Fact II: Gains & Losses Among Sector Switches

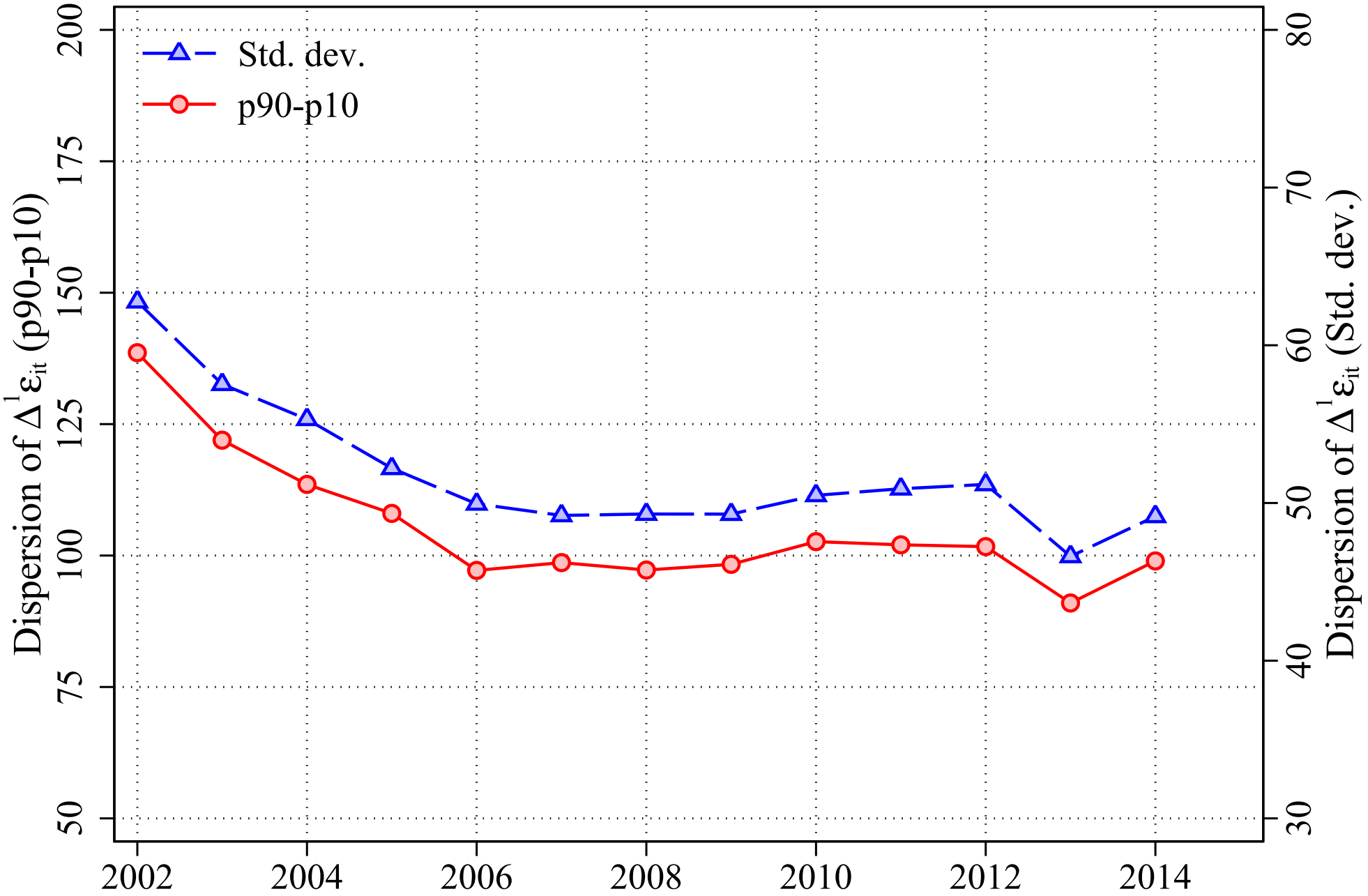


Fact III: Comparison with High- vs. Low-Income Dynamics in US

	Std. dev.	Skewness	Kurtosis
Panel A. Brazil PME			
All of Brazil	0.52	-0.04	7.17
Formal-Formal	0.37	-0.17	9.61
Informal-Informal	0.60	0.00	5.40
Formal-Informal	0.86	-0.10	3.50
Informal-Formal	0.78	0.15	3.67
Panel B. Comparison with US SSA (Guvenen et al. '19)			
Average income (P50)	0.49	-1.35	16.81
High income (P90)	0.43	-1.62	26.20
Low income (P10)	0.73	-0.72	6.78

Note: US statistics are from Guvenen et al. ('19) for men in age group 3, see data file tab "L1_log_age_re".

Fact IV: Decline in Income Volatility in Overall Economy

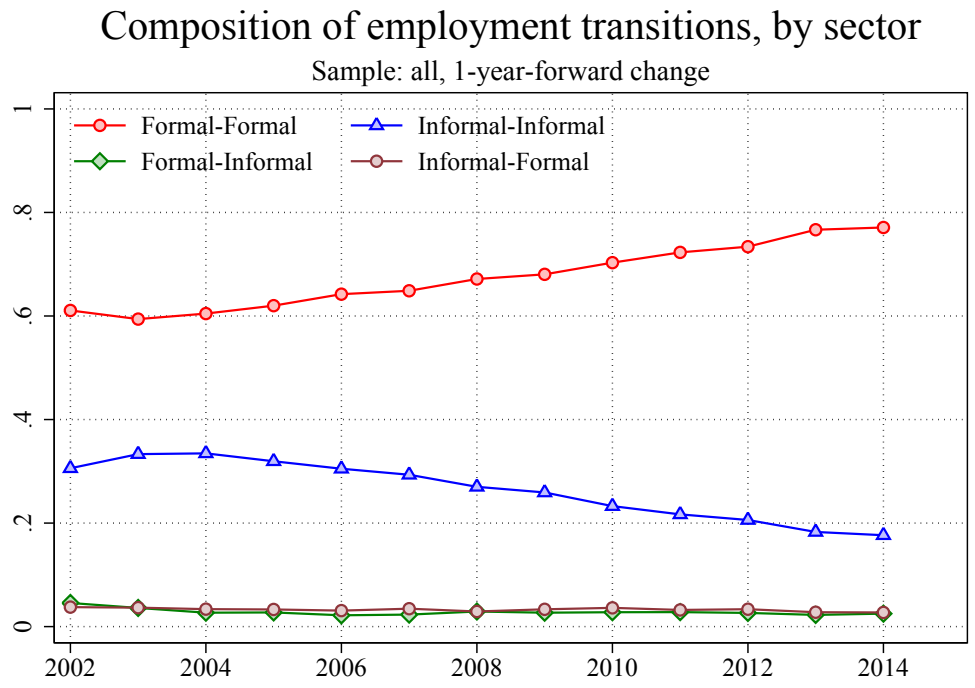


Fact V: Sectoral & Within-Sector Changes

▸ By age ▸ By education

Decomposition of (within-sector) decline in income volatility:

1. Shift toward formal sector—accounts for app. 40%
2. Changes within sectors—account for app. 60%



	St.d.$(\Delta^1 \varepsilon_{it})$	
	2002	2014
All	0.63	0.49
F-F	0.44	0.39
F-I	0.97	0.81
I-F	0.79	0.88
I-I	0.80	0.63

Conclusion

Main Take-Aways & Future Work

- Large decline in income inequality and volatility in Brazil
- Room for future work to explore deeper determinants:
  - How much of time trends is due to compositional changes?
  - Are formal and informal sector workers really different?
  - Can earnings dynamics be linked to firm heterogeneity?
  - Role of policies, e.g. rapid rise in minimum wage?
  - Higher-frequency earnings fluctuations, e.g. monthly?
  - Effects of Brazil's recent Great Recession 2014–2016?

...thank you for your questions & comments!

Appendix

Backup: Data Compatibility ▸ Back

Comparing admin. data (RAIS) & household survey (PME):

- Restrict RAIS to 6 metro regions covered by PME 2002–15.

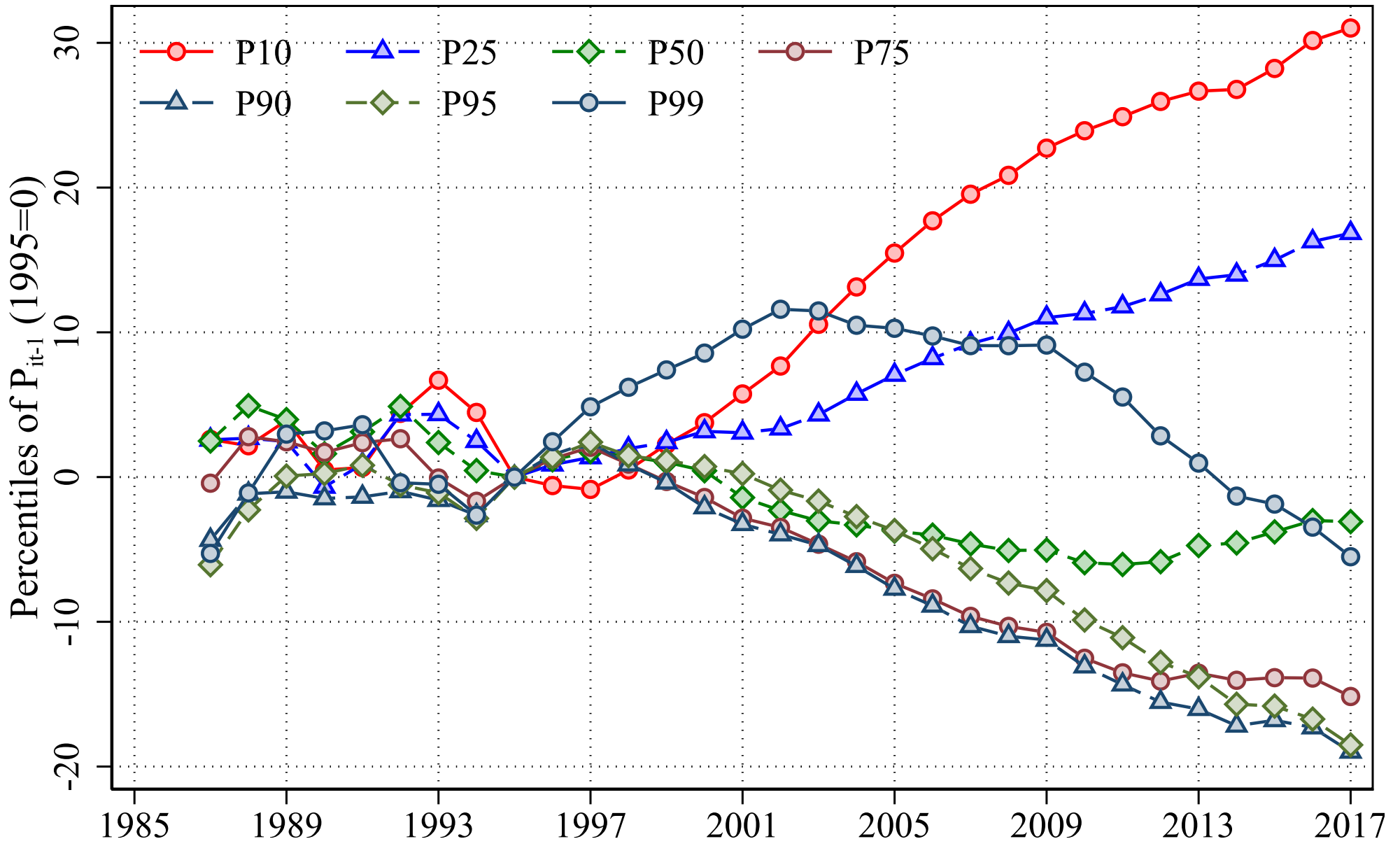
Potential discrepancies:

- Coverage: RAIS is employees only, PME also contractors.
- Sample size: RAIS captures far right tail, PME less so.
- Timing: PME is rolling 4-month window, RAIS all-year.
- Reporting: RAIS is employer-reported, PME self-reported.
- Response: RAIS is mandatory, PME voluntary.
- Coding: RAIS is top-coded at 120 (150) $\times$MW, PME not.
- Accuracy: RAIS is cross-checked and audited, PME not.

Conclusion: Good cross-sectional match, diverging dynamics.

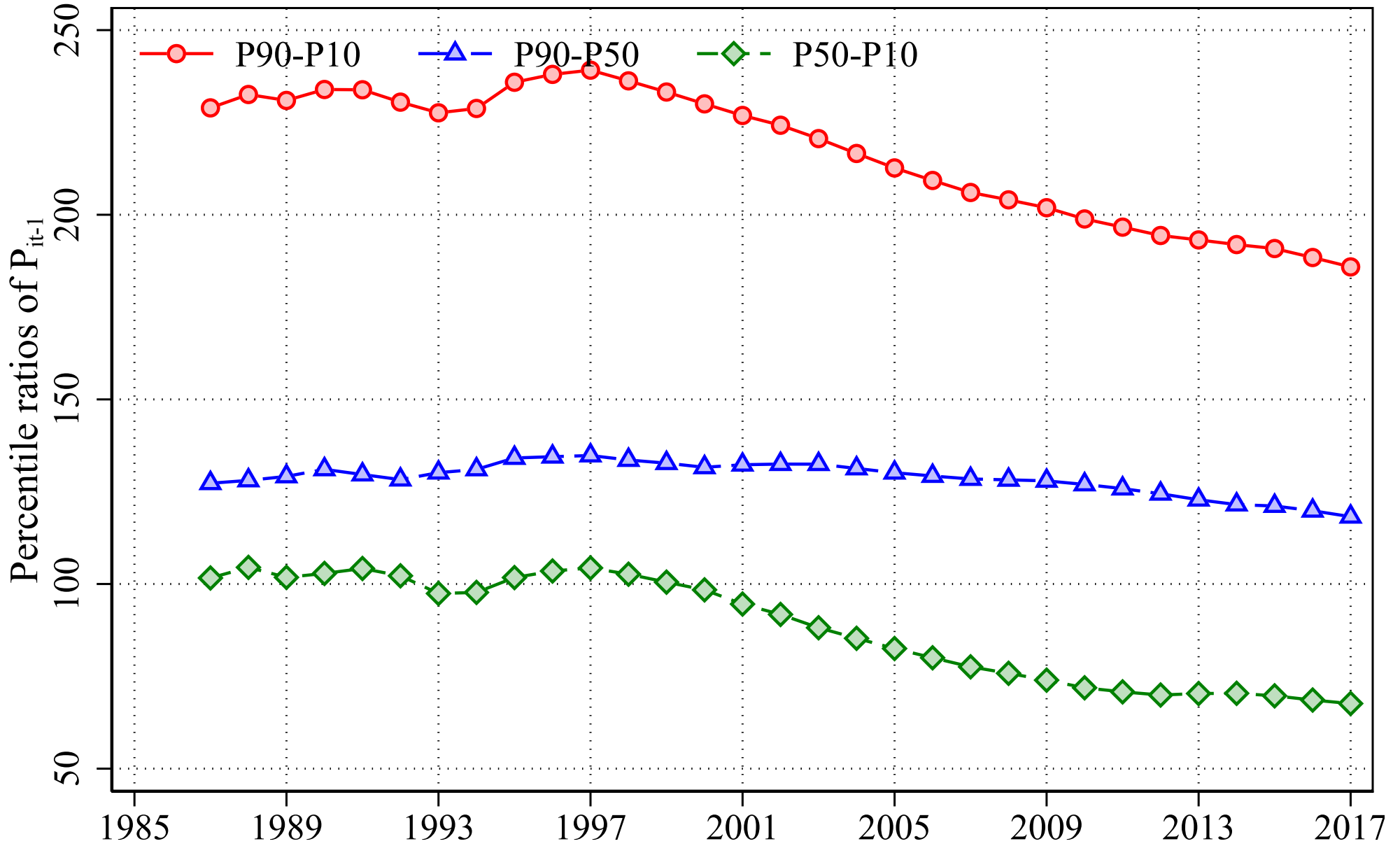
Backup: Norm. Log Permanent Earnings Percentiles

▸ Back



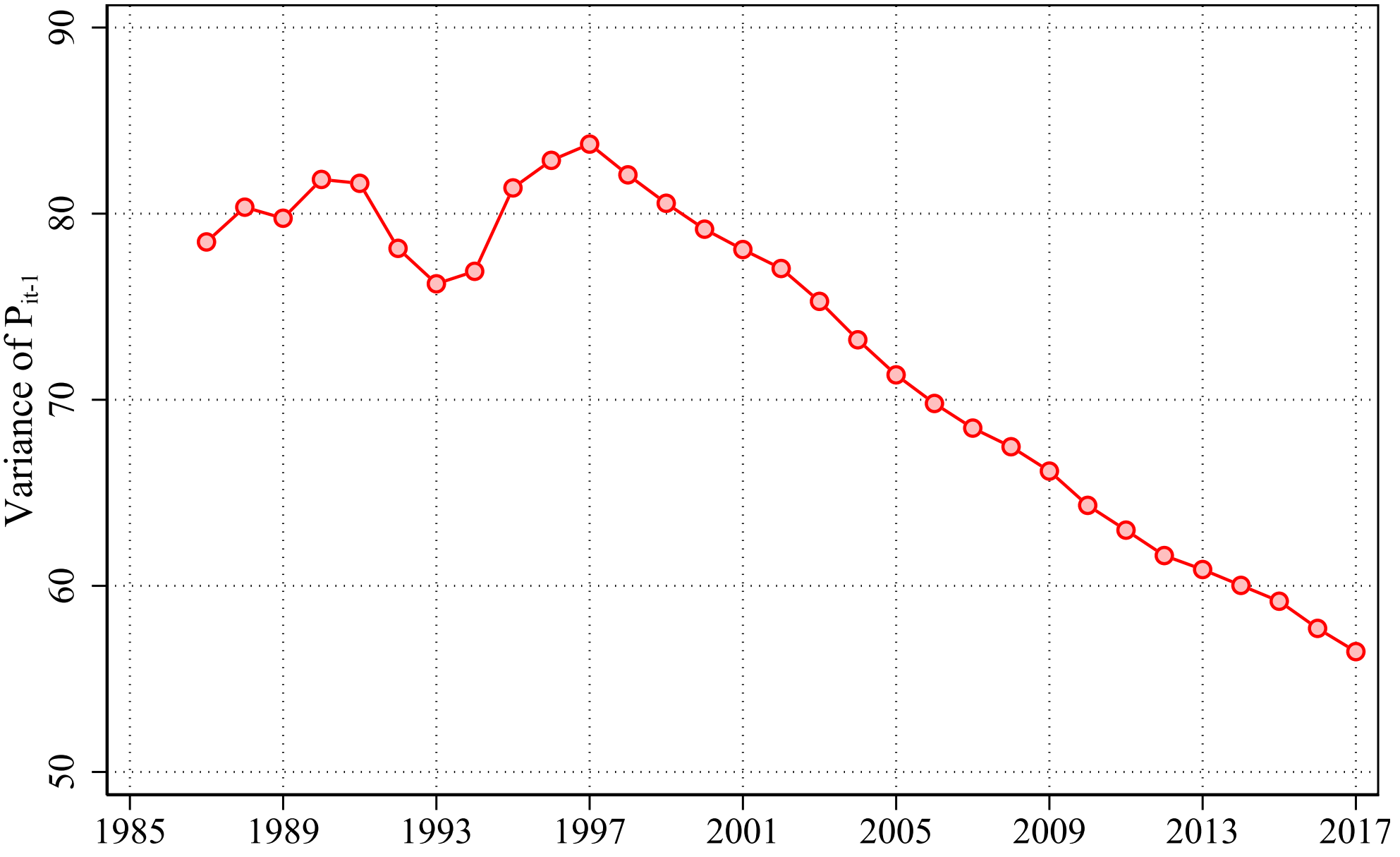
Backup: Log Permanent Earnings Percentile Ratios

▸ Back



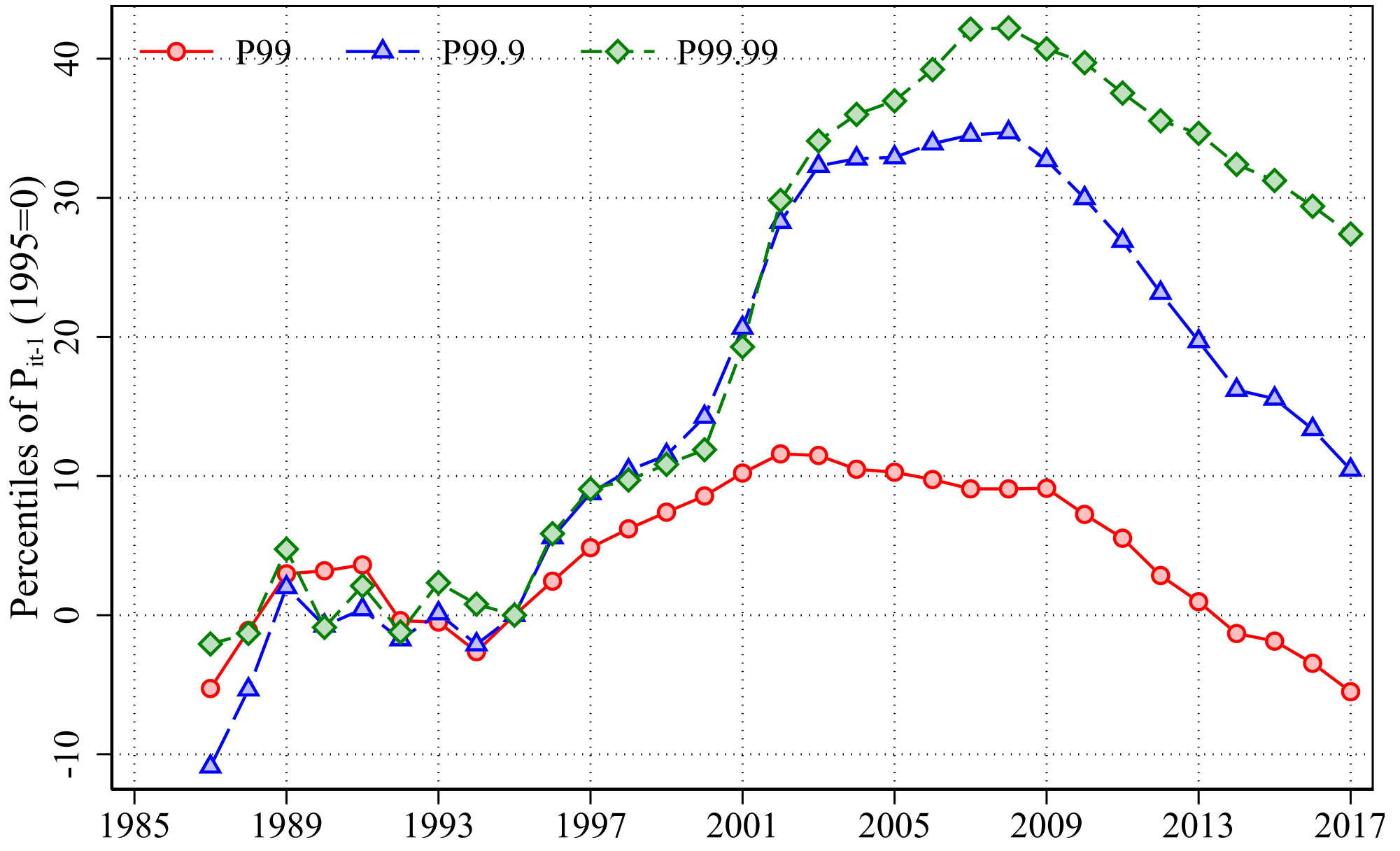
Backup: Variance of Log Permanent Earnings

▸ Back



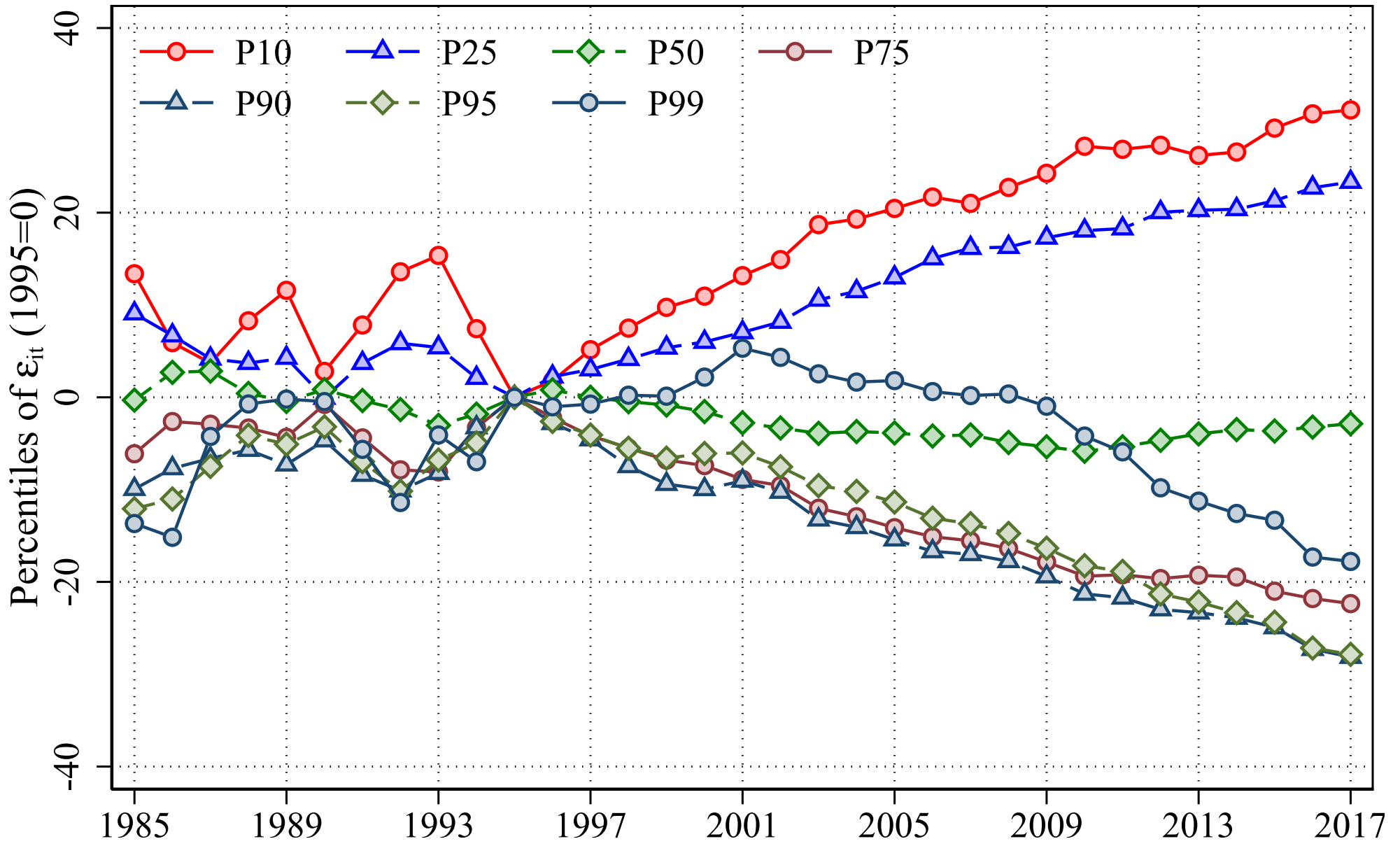
Backup: Top Log Permanent Earnings Percentiles

▸ Back



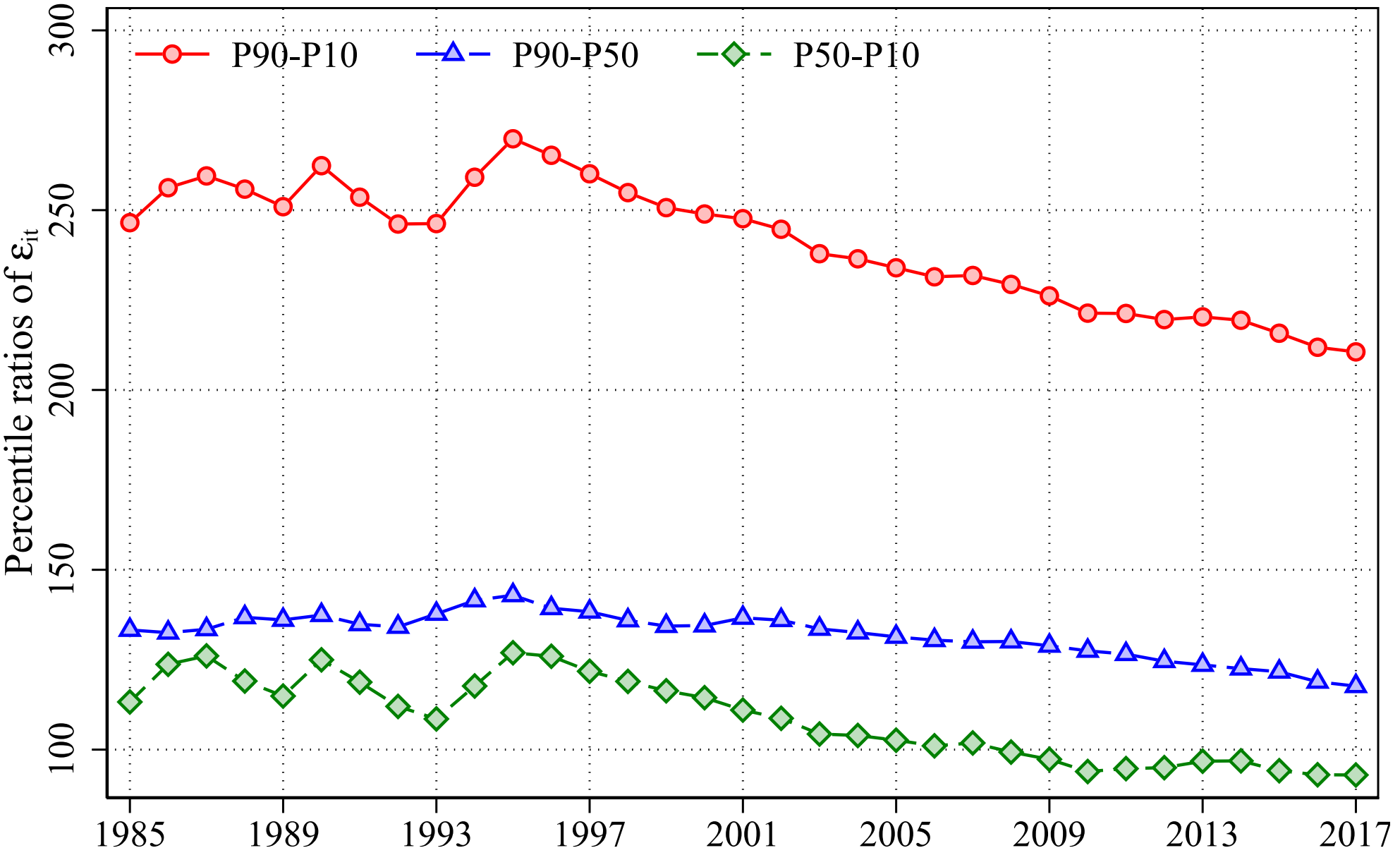
Backup: Normalized Log Residual Earnings Percentiles

▸ Back



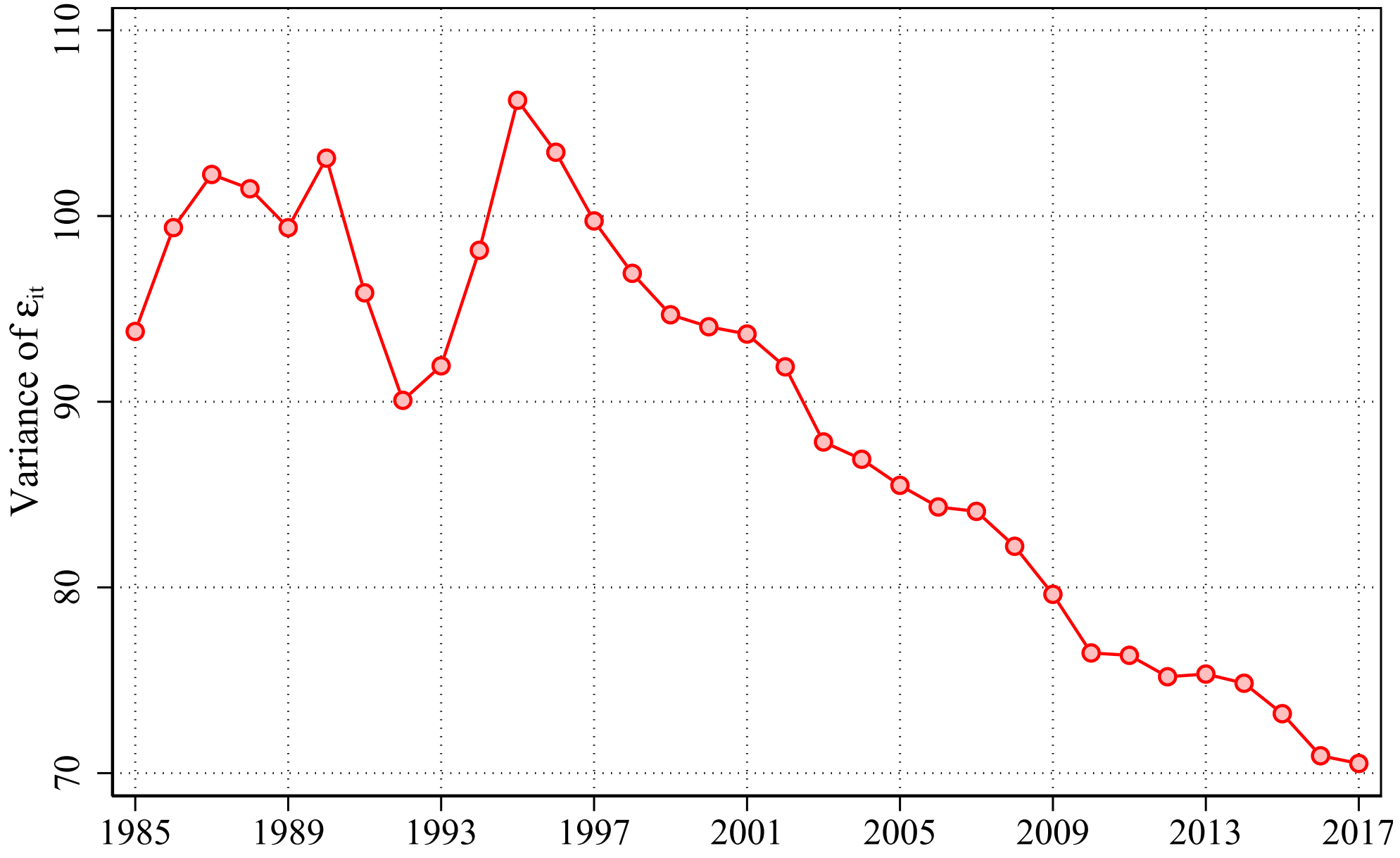
Backup: Log Residual Earnings Percentile Ratios

▸ Back



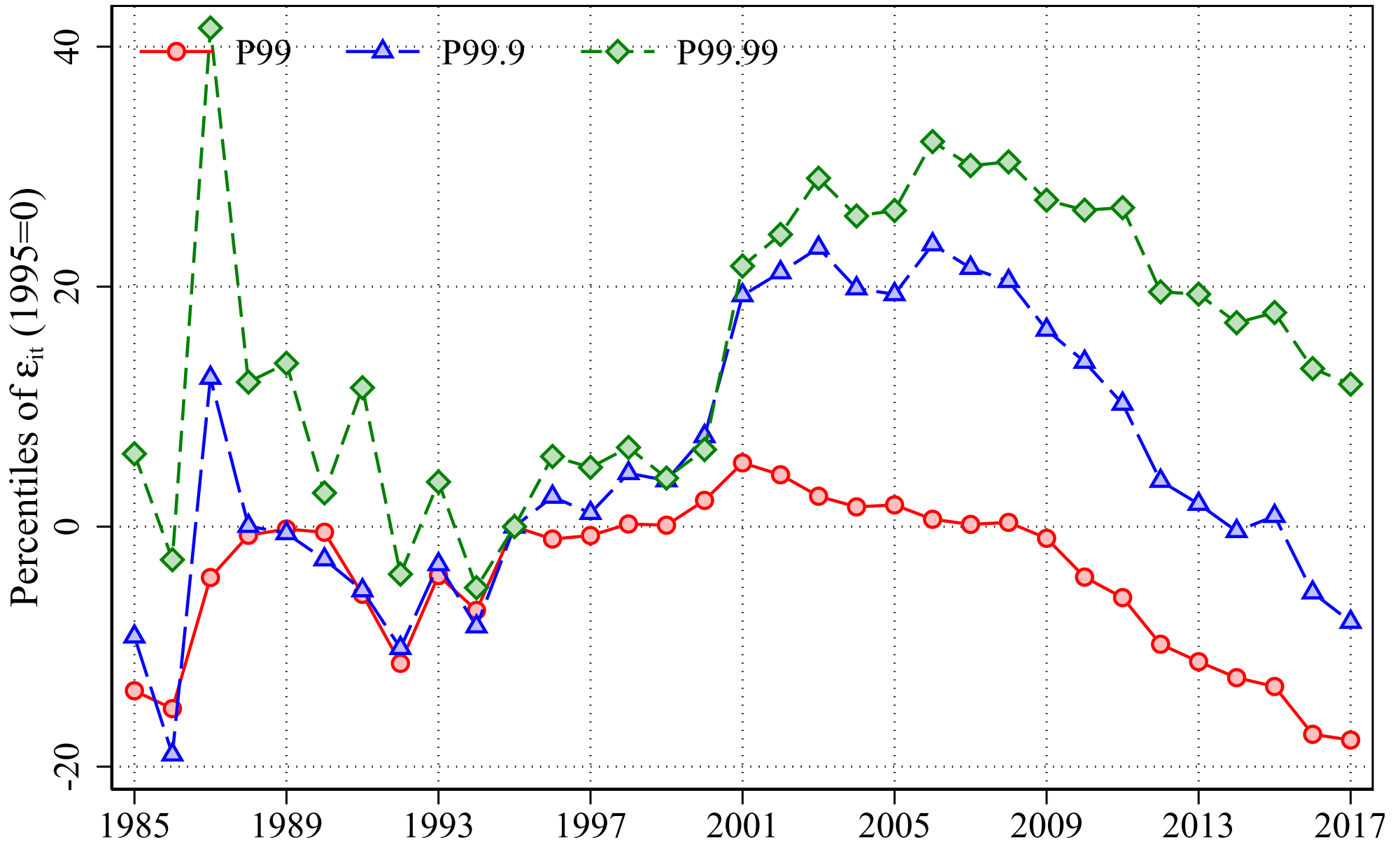
Backup: Variance of Log Residual Earnings

▸ Back



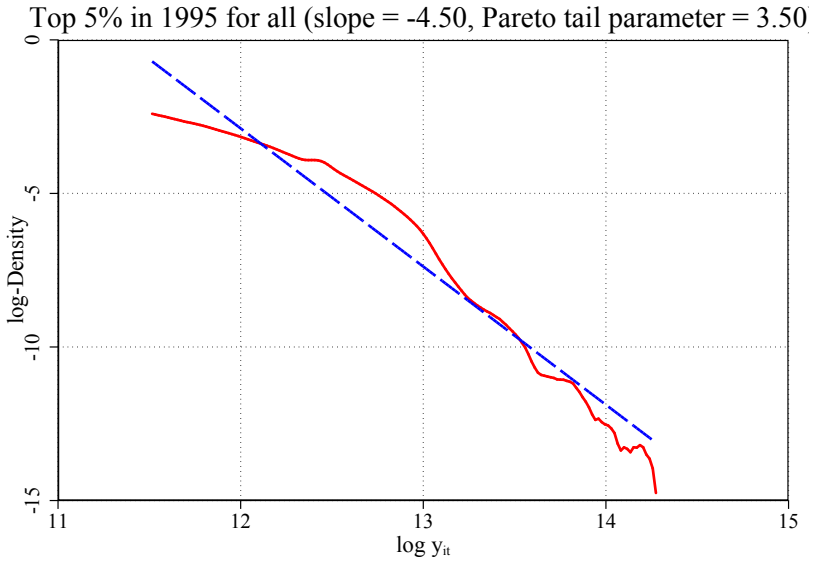
Backup: Top-Log-Residual Earnings Percentiles

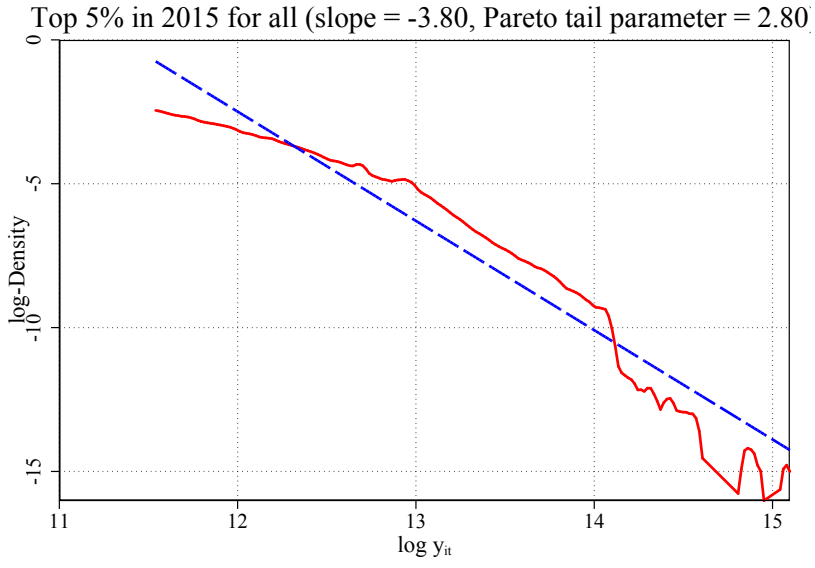
▸ Back

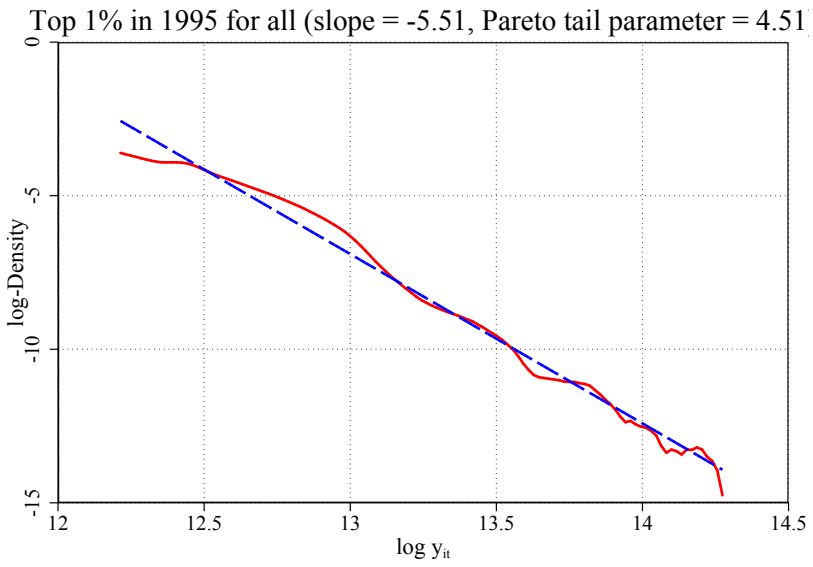


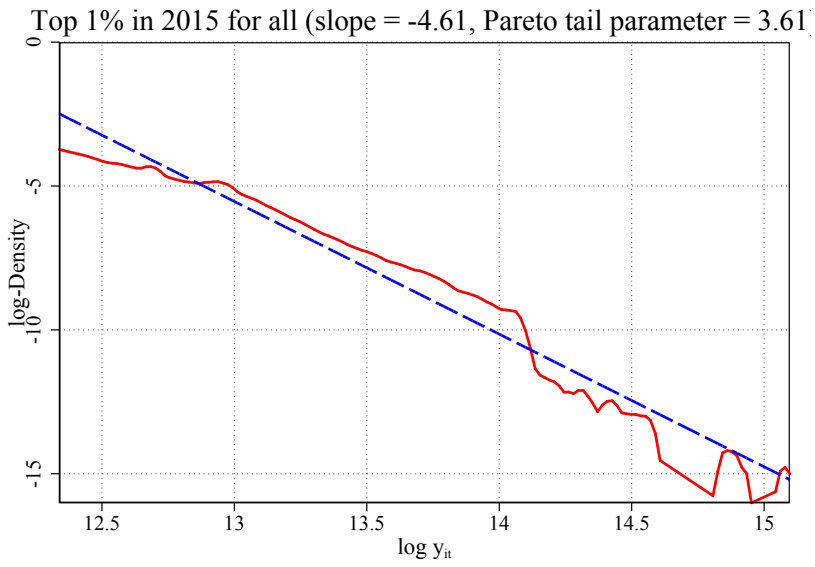
Backup: Pareto Tail of Earnings, Log-Log Dens. Plots

▸ BACK







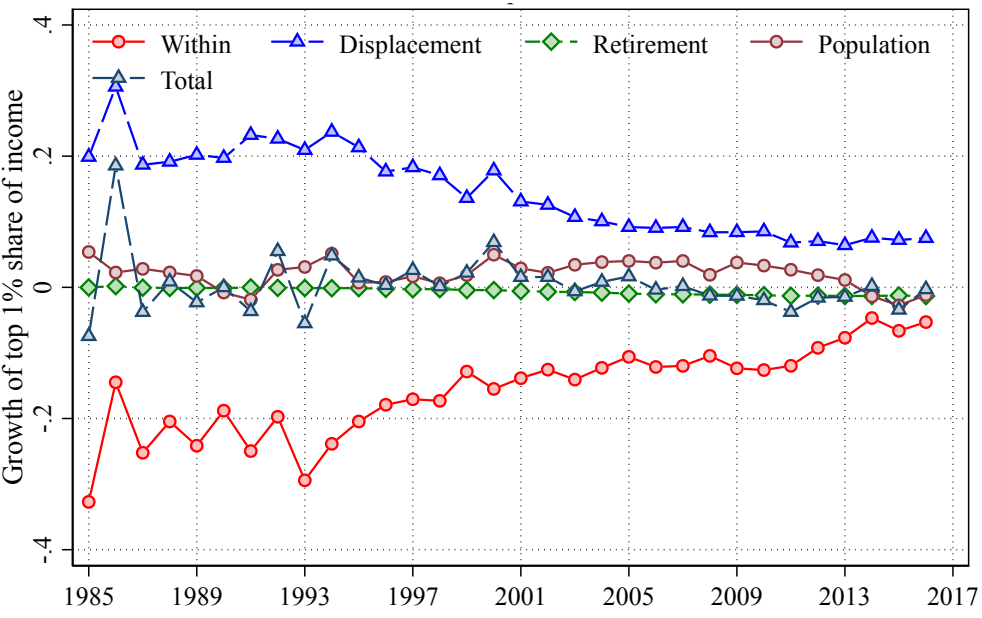


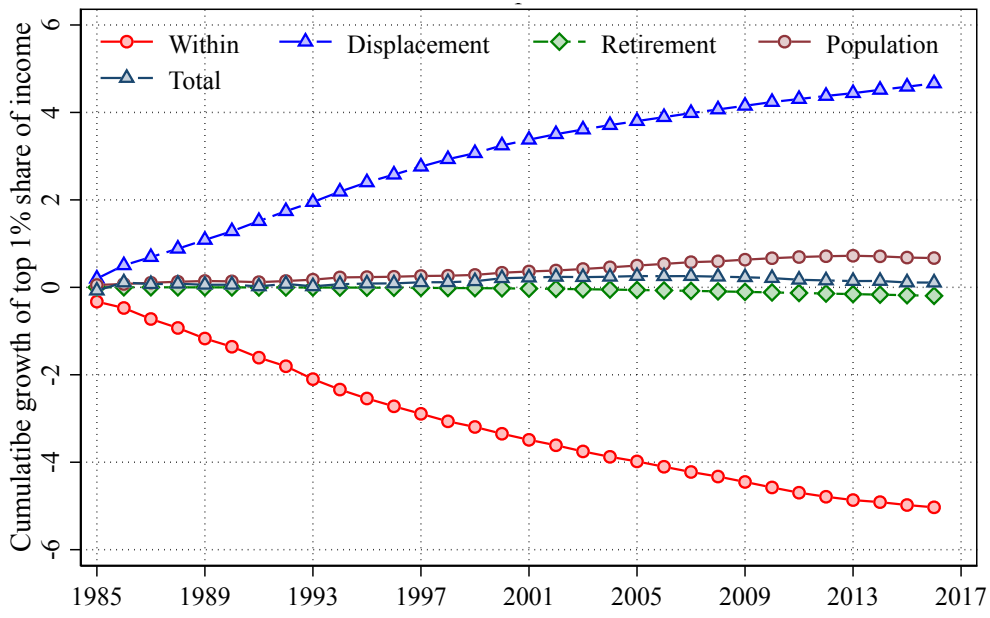
Note: Pareto tail $\alpha = (-\text{slope} - 1)$ in log-log density plots.

Backup: Gomez ('19) Decomp. of Top Income Growth

▸ Back

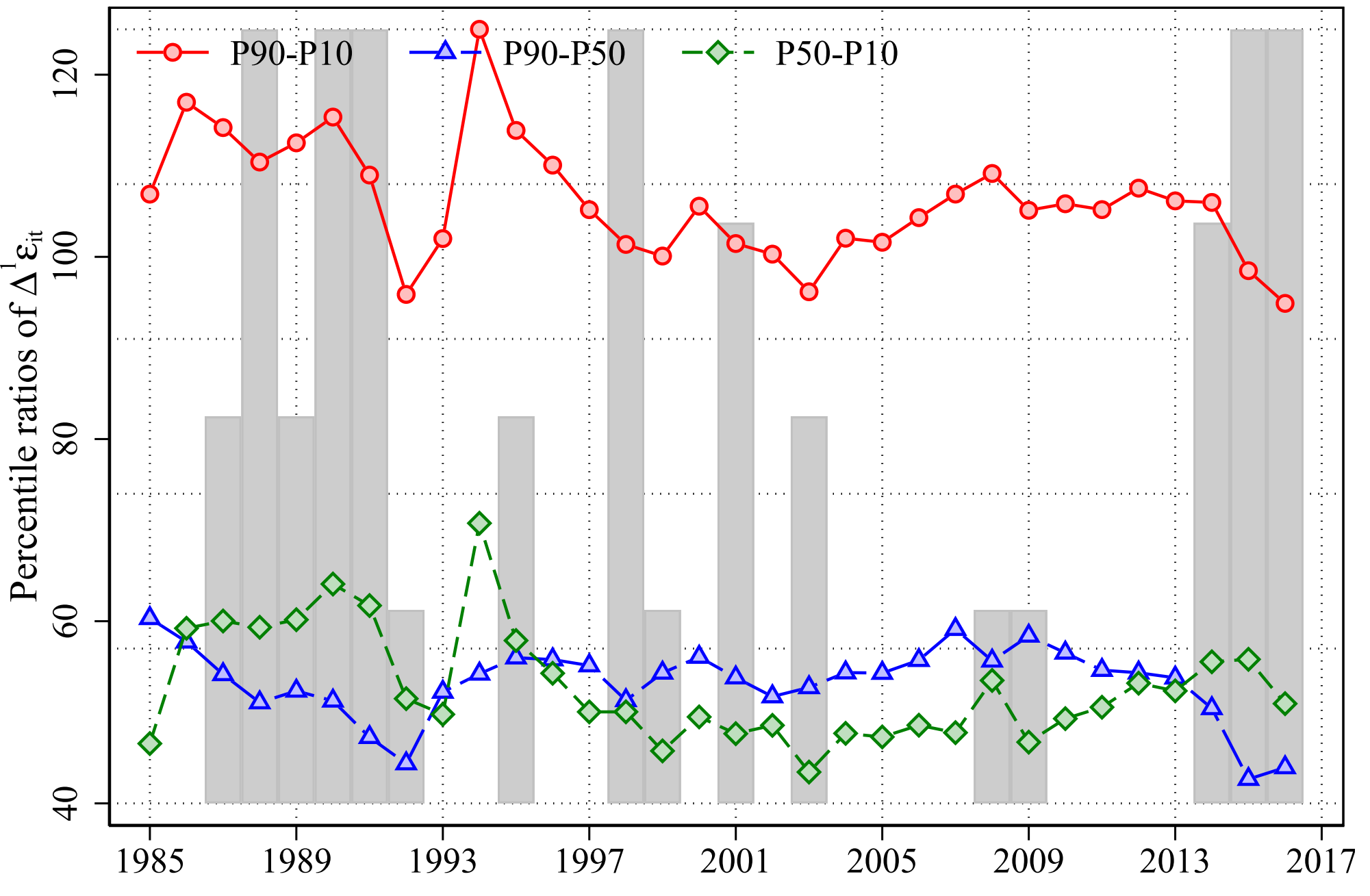
- Left: positive displacement-term, negative within-term, both declining in magnitude over tme
- Right: cumulatively, positive displacement-term and negative within-term strongly dominate over time





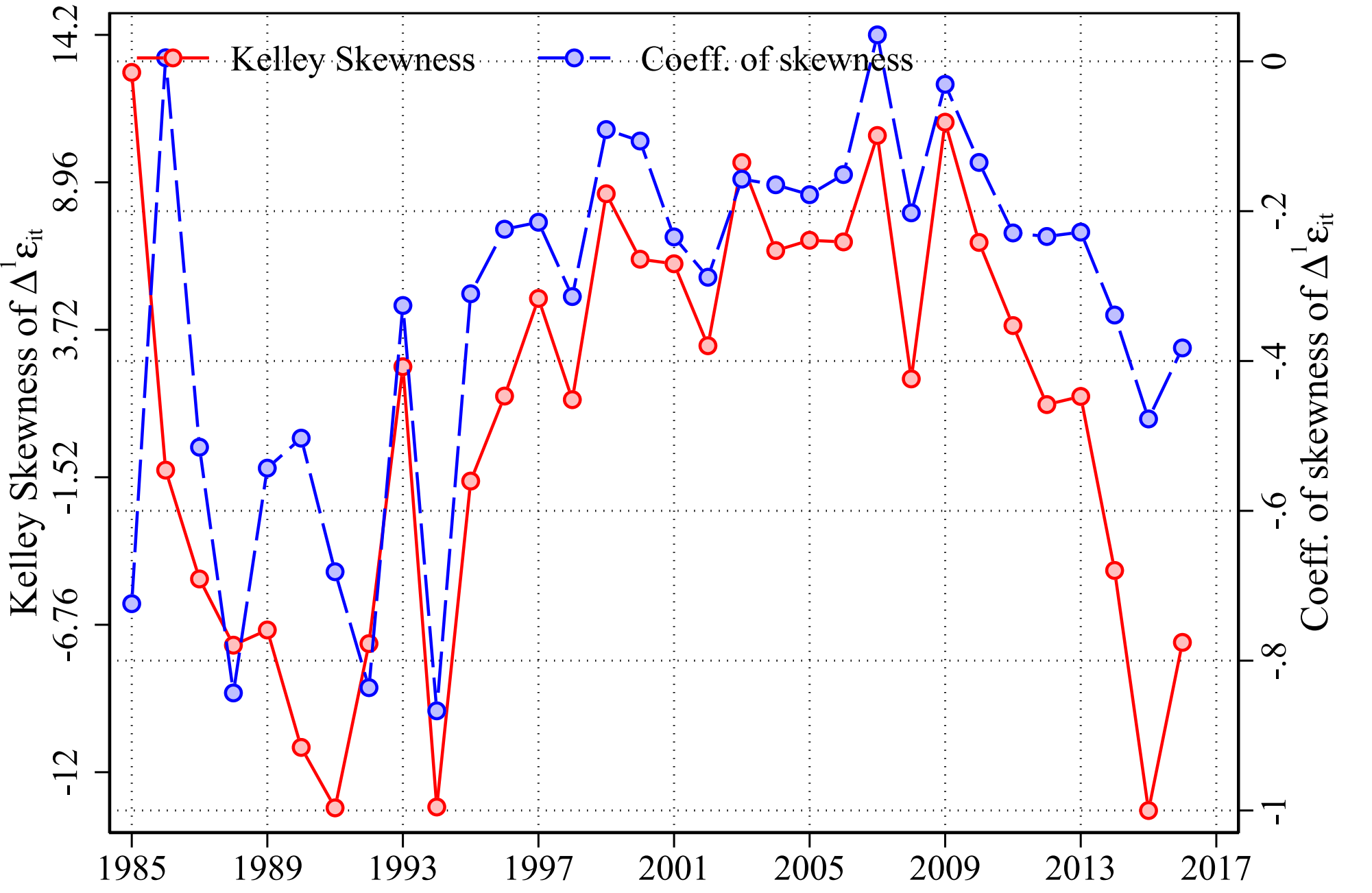
Backup: Log Perc. Ratios of 1-Year Resid. Earnings Ch.

▸ Back



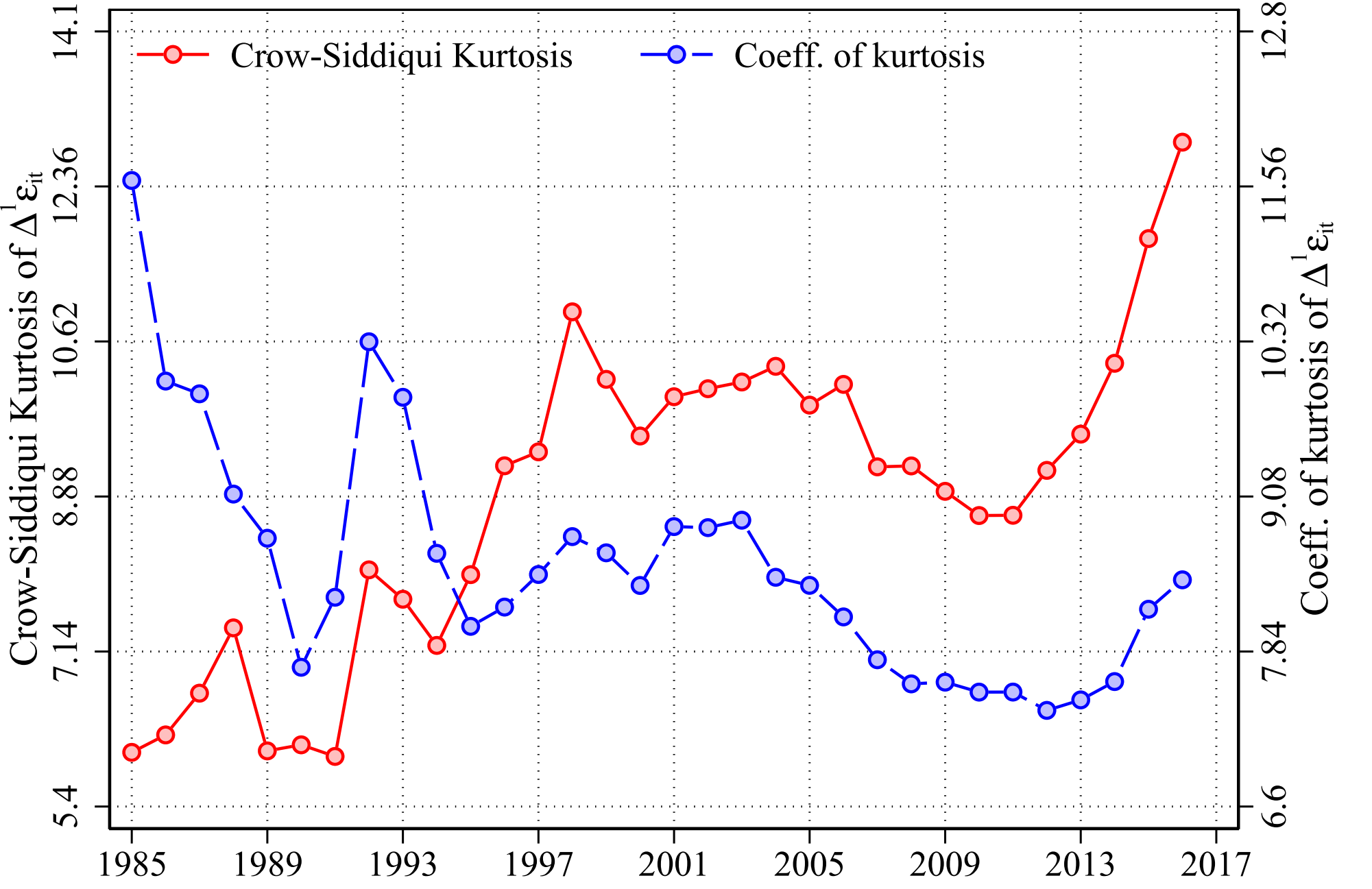
Backup: Skewness of 1-Year Resid. Earnings Change

▸ Back



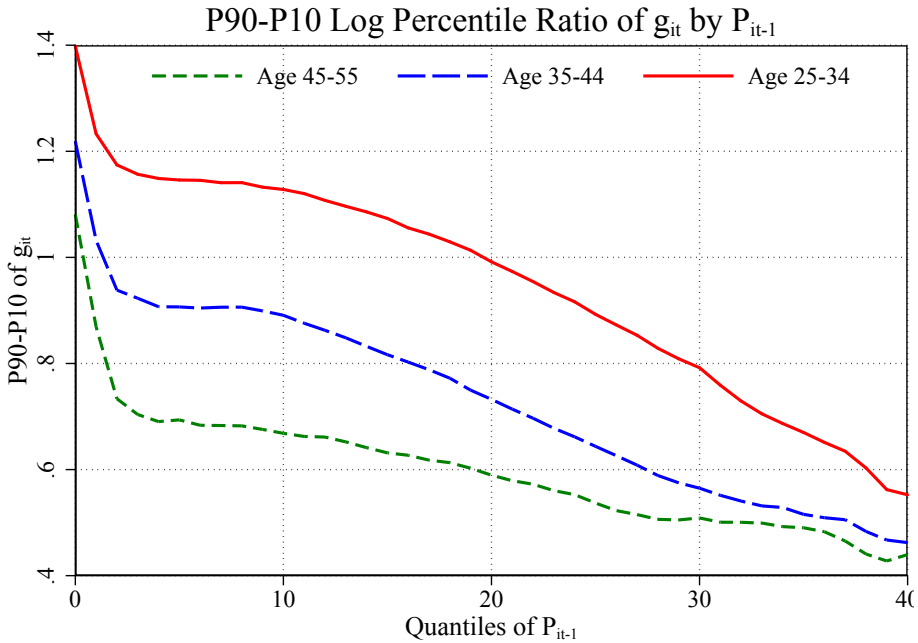
Backup: Kurtosis of 1-Year Resid. Earnings Change

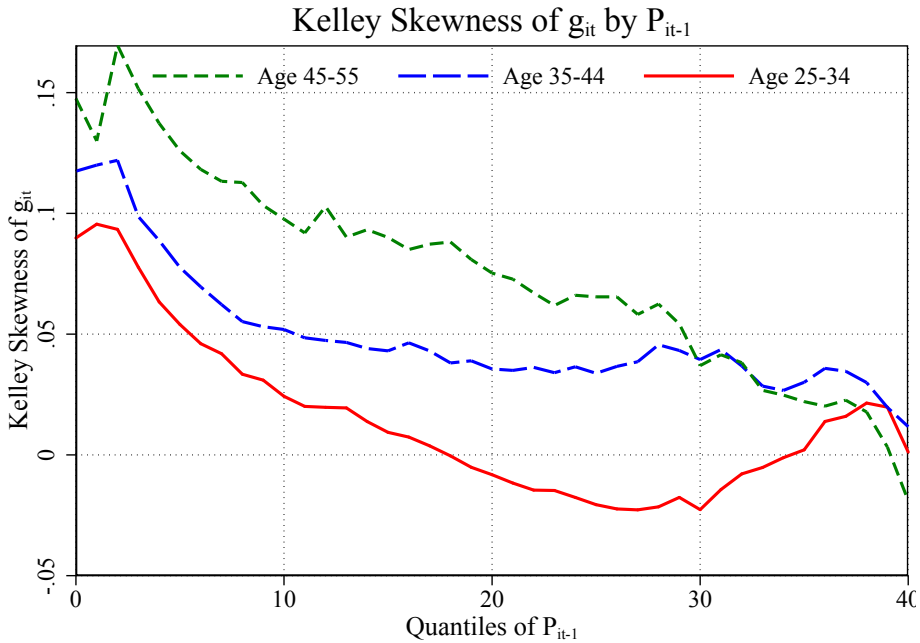
▸ Back

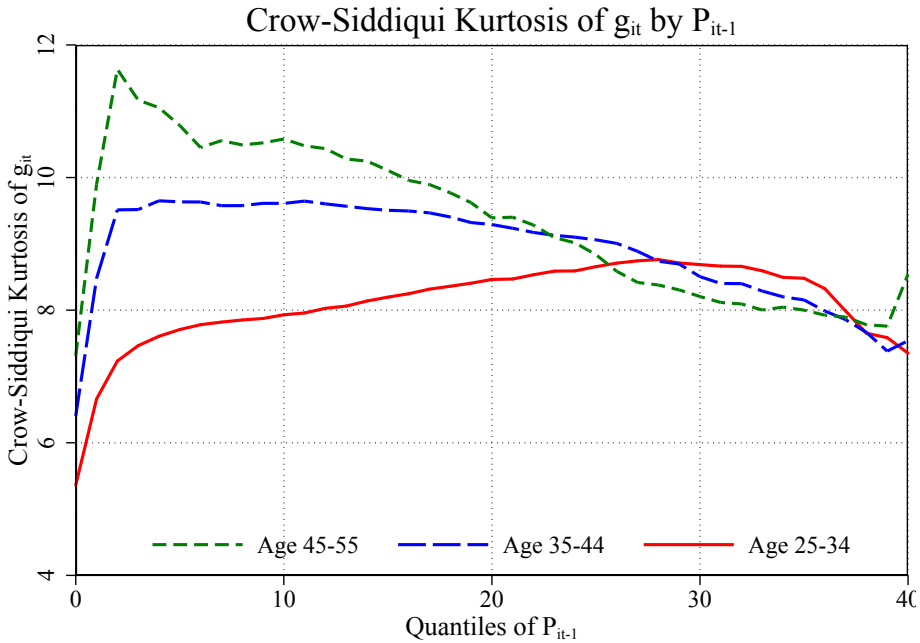


Backup: 1-Year Resid. Earnings Ch., by Age & Perm. Inc.

▸ BACK

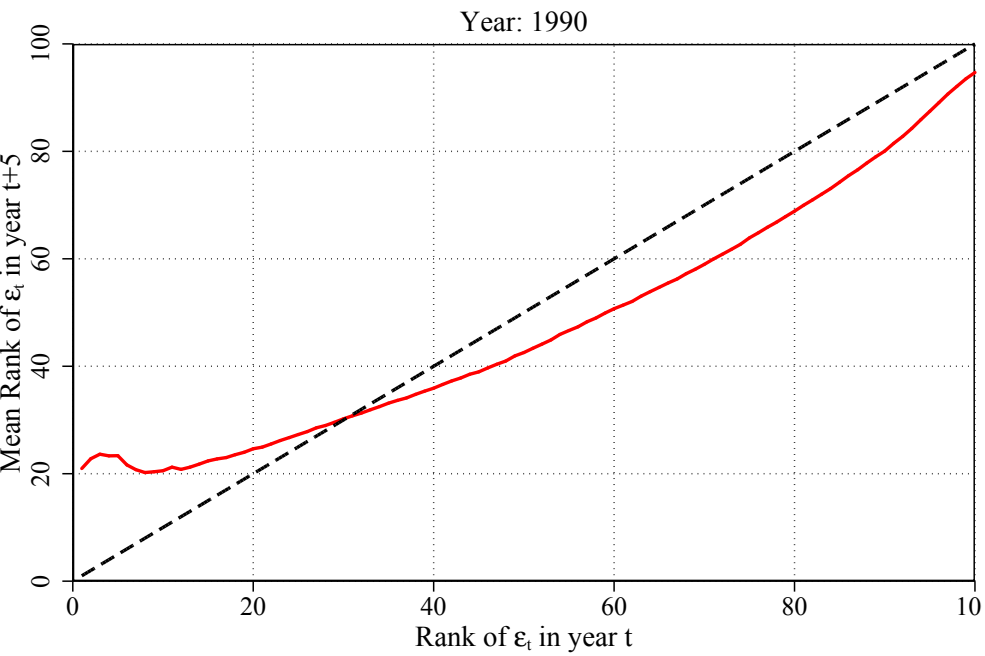


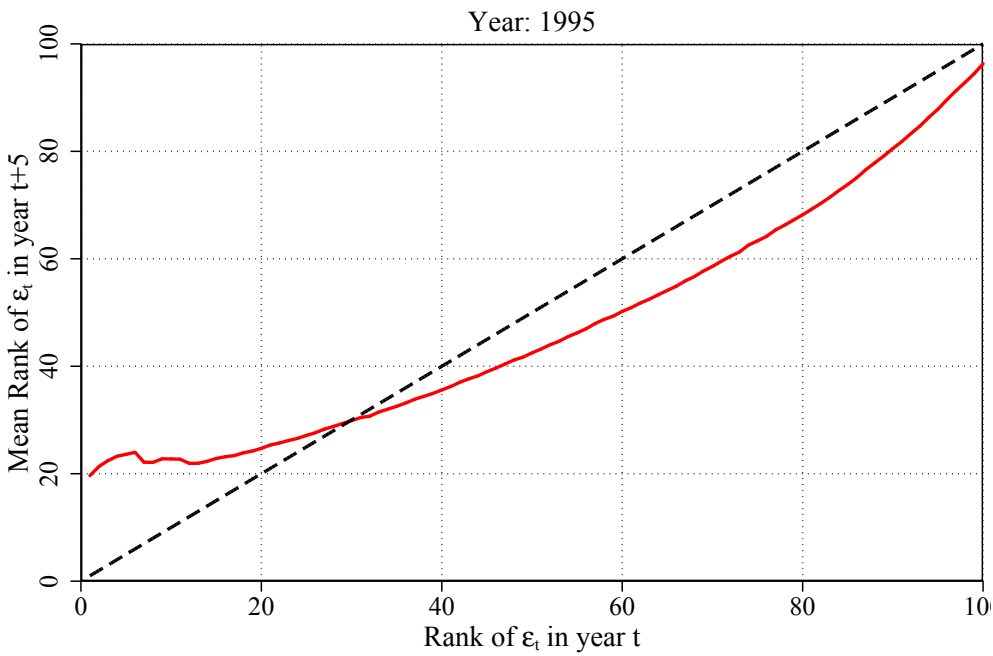


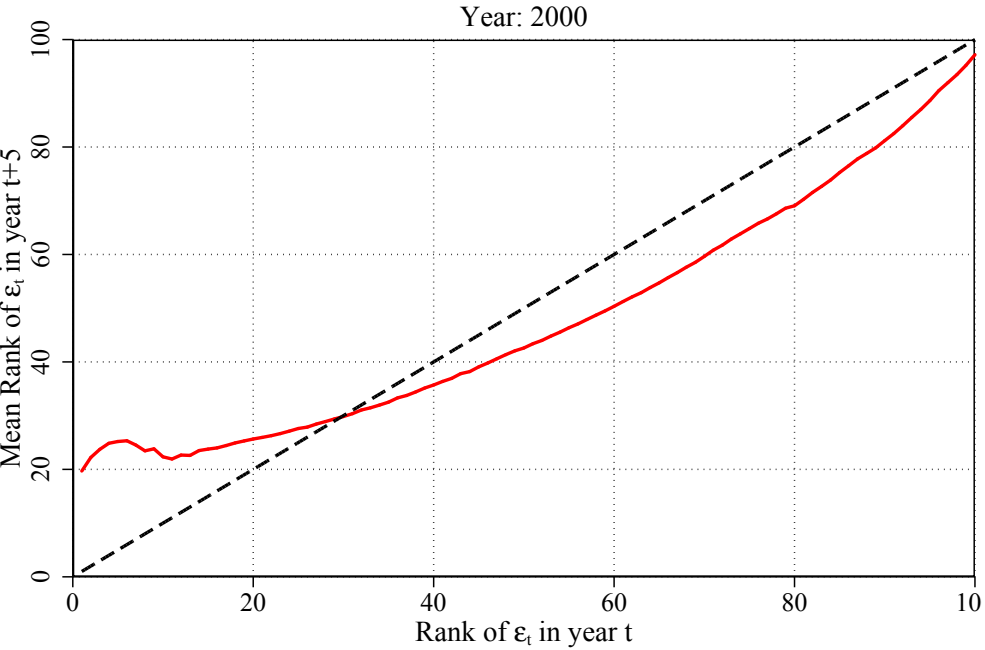


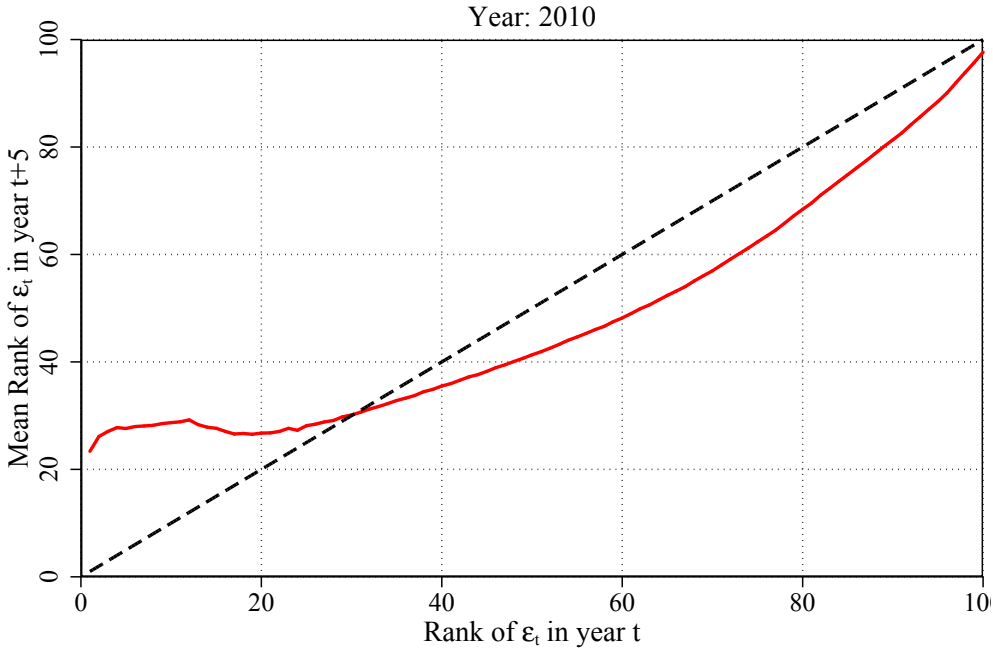
Backup: Long-Term Mobility in Resid. Inc., Various Years

▸ Back



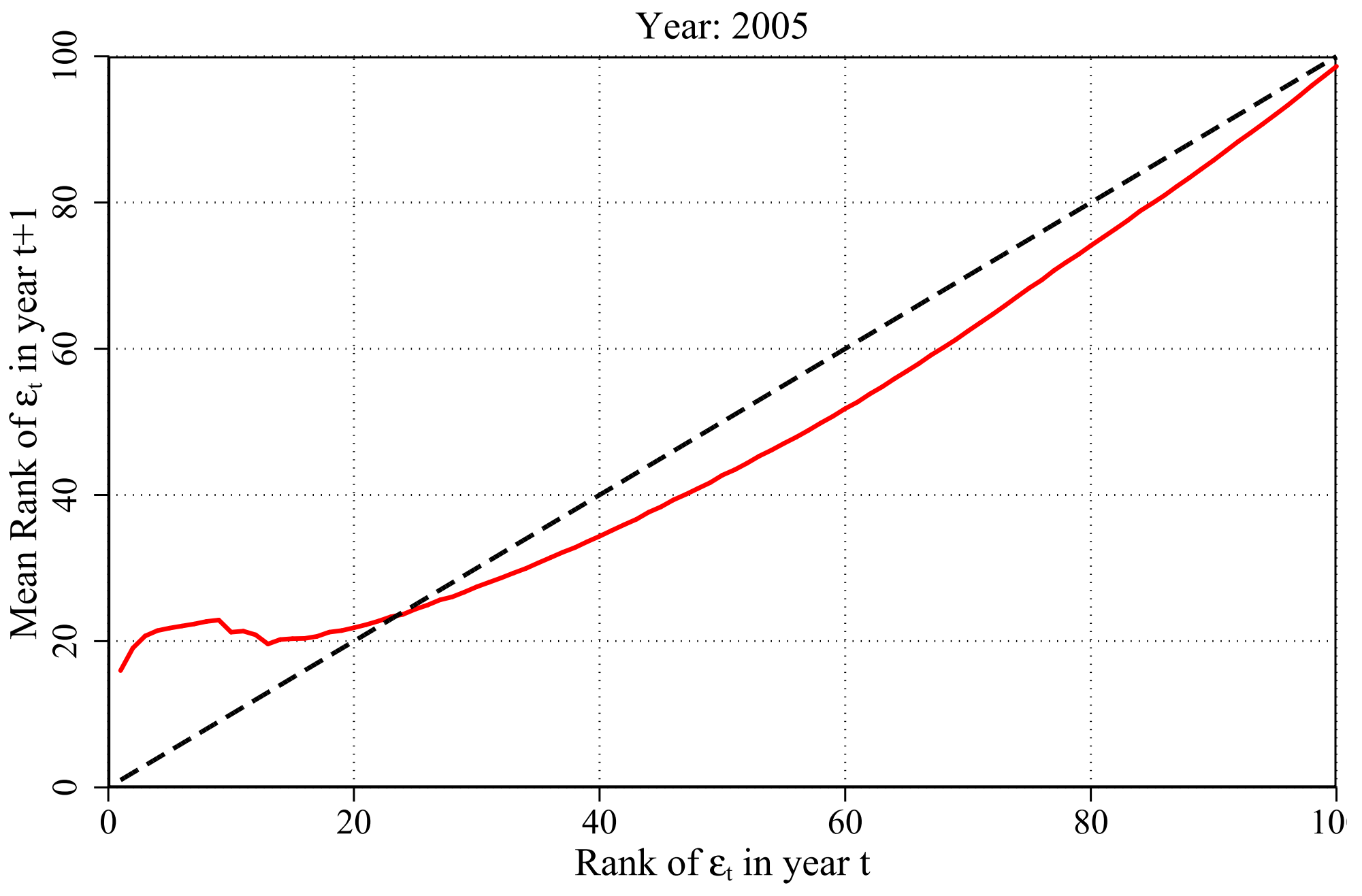






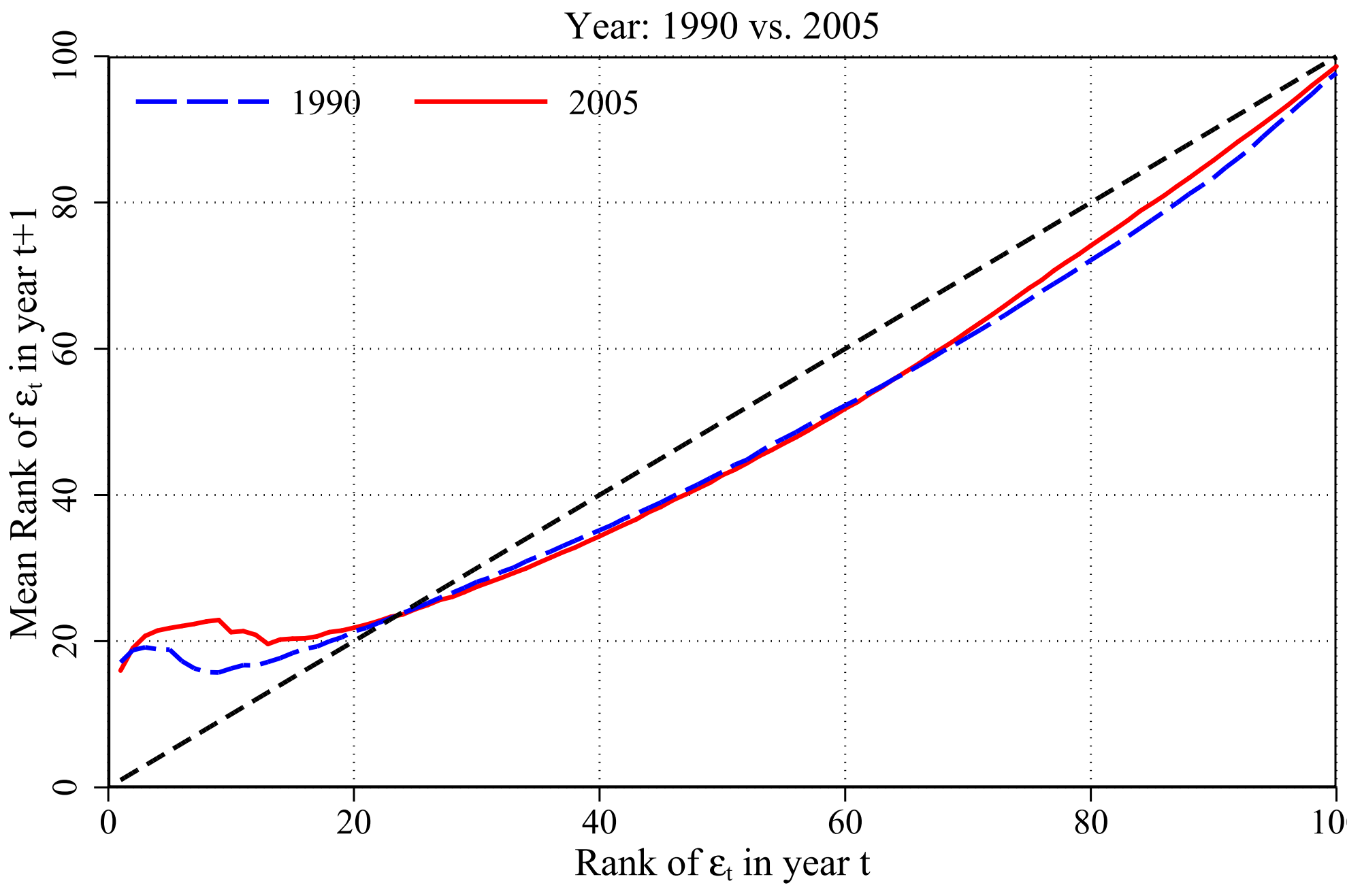
Backup: Short-Term Mobility in Resid. Income, 2005

▸ Back



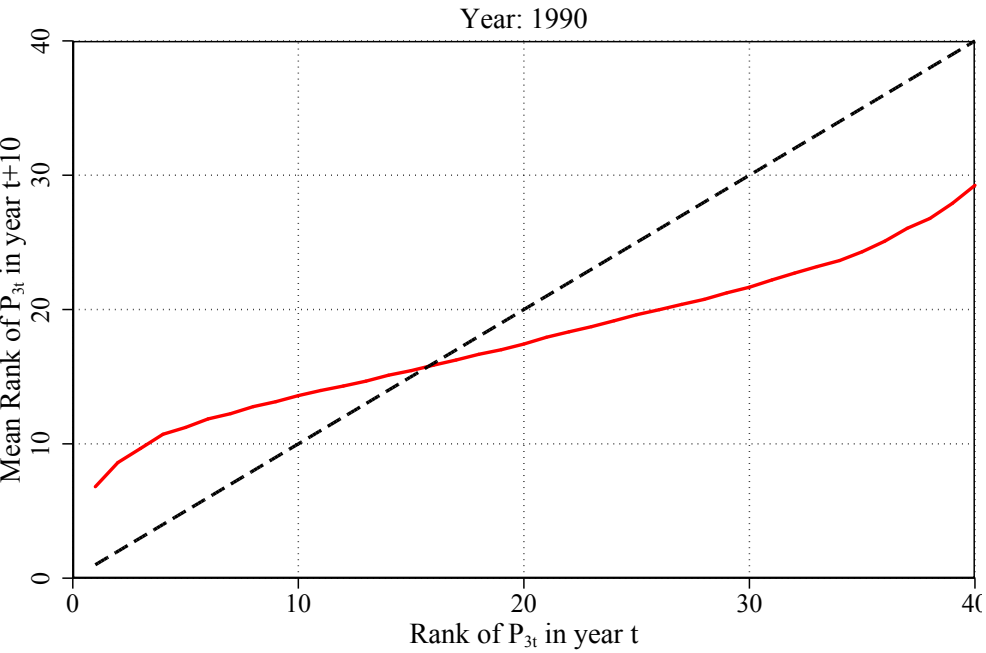
Backup: Short-Term Mobility in Resid. Inc., 1990 vs. 2005

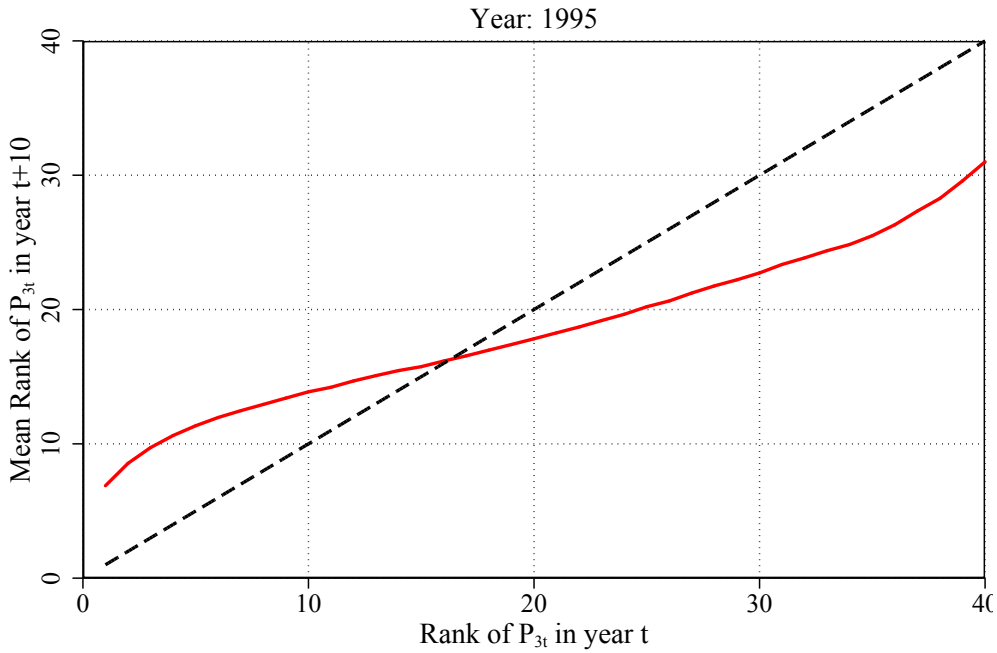
▸ Back

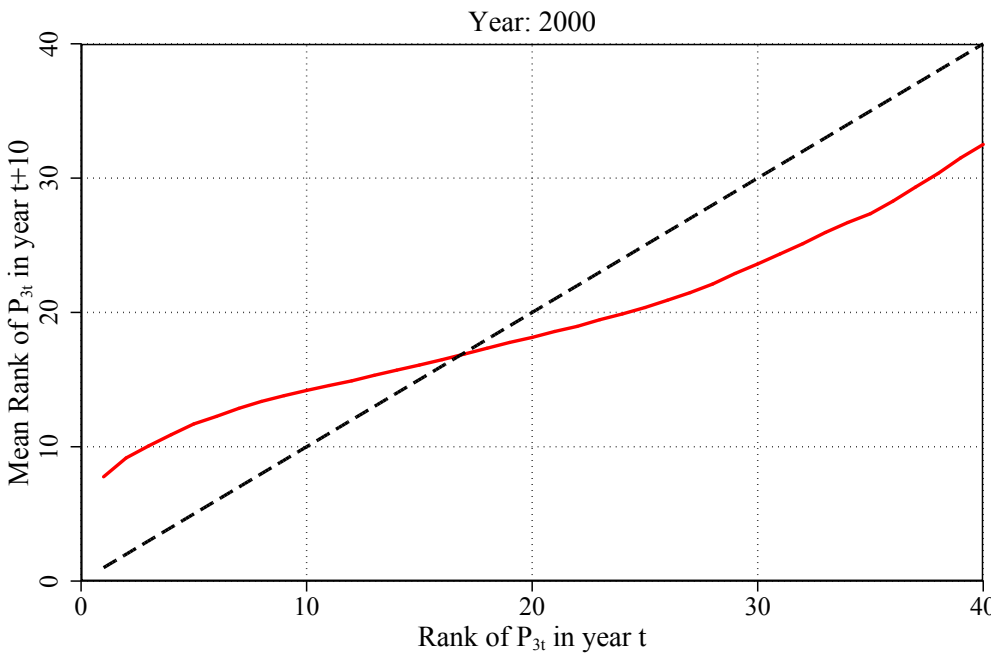


Backup: Perm Mobility in Resid. Income, Various Years

▸ Back

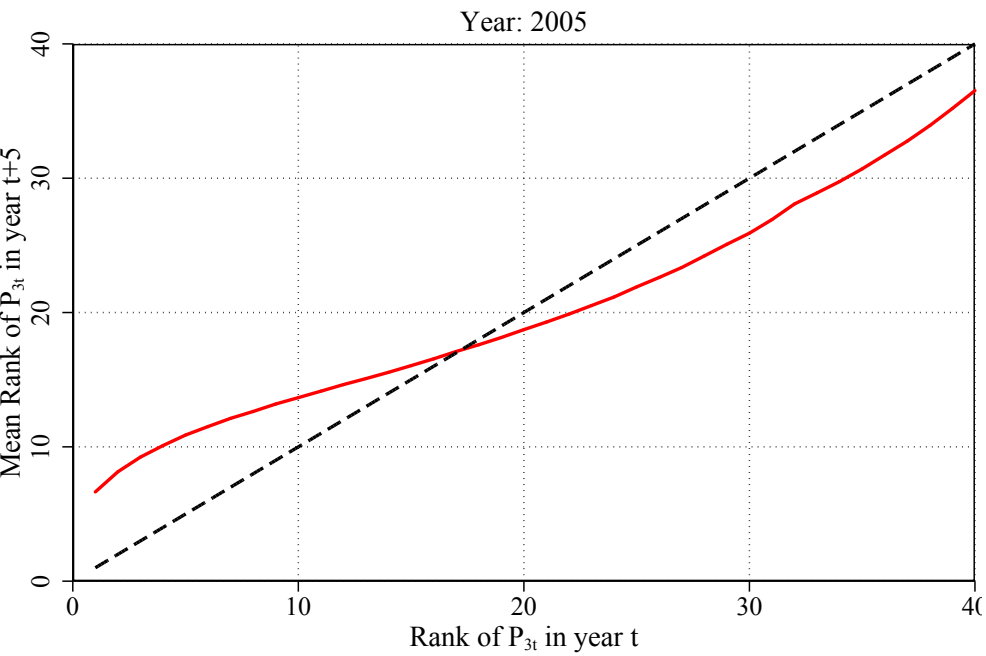


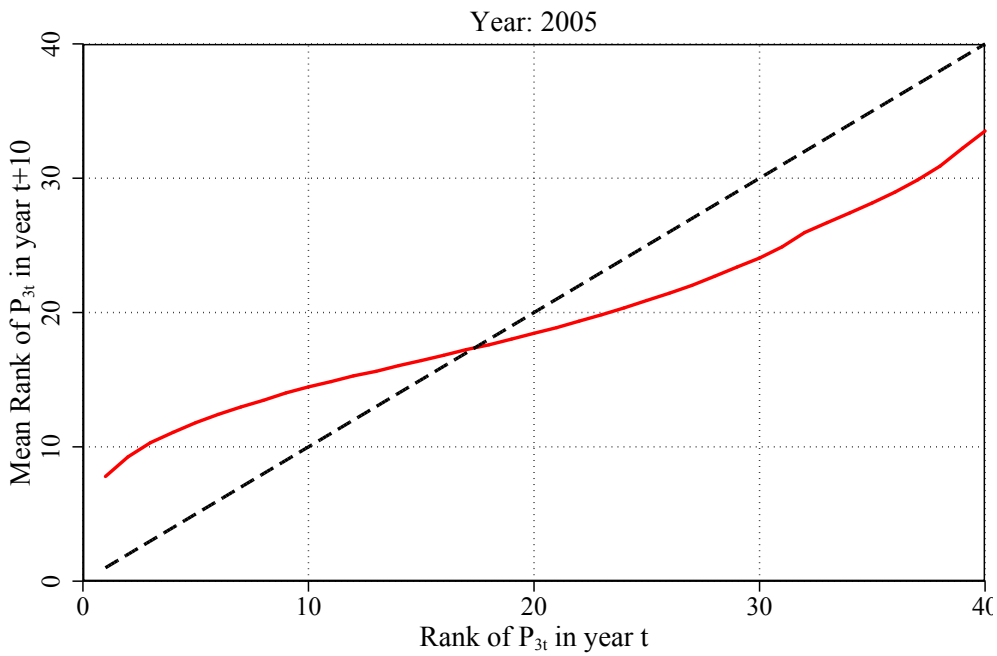


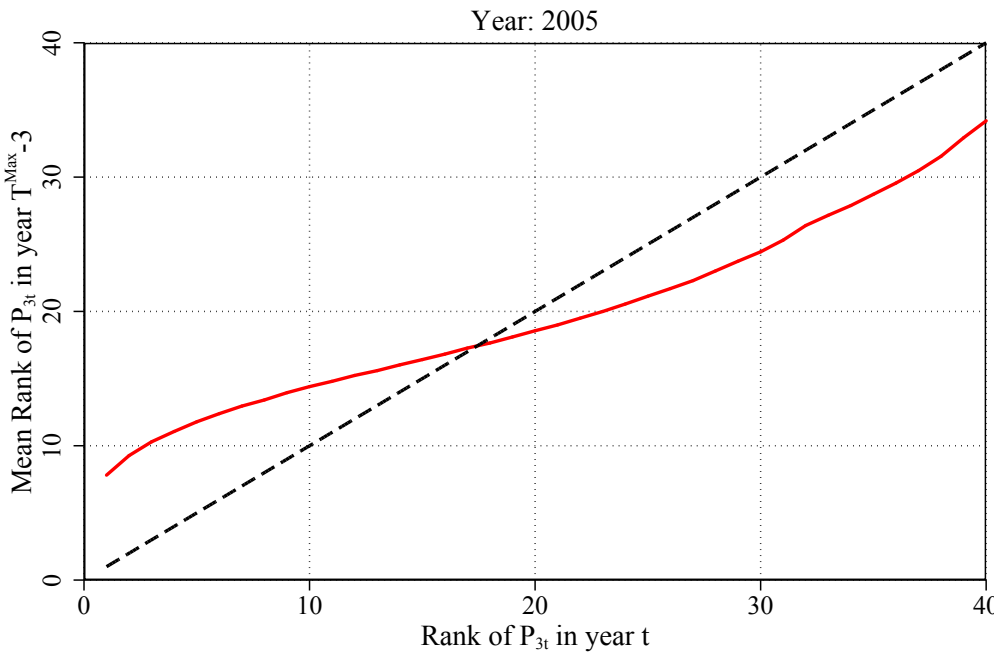


Backup: Perm. Mobility in Residual Income, 2005

▸ Back

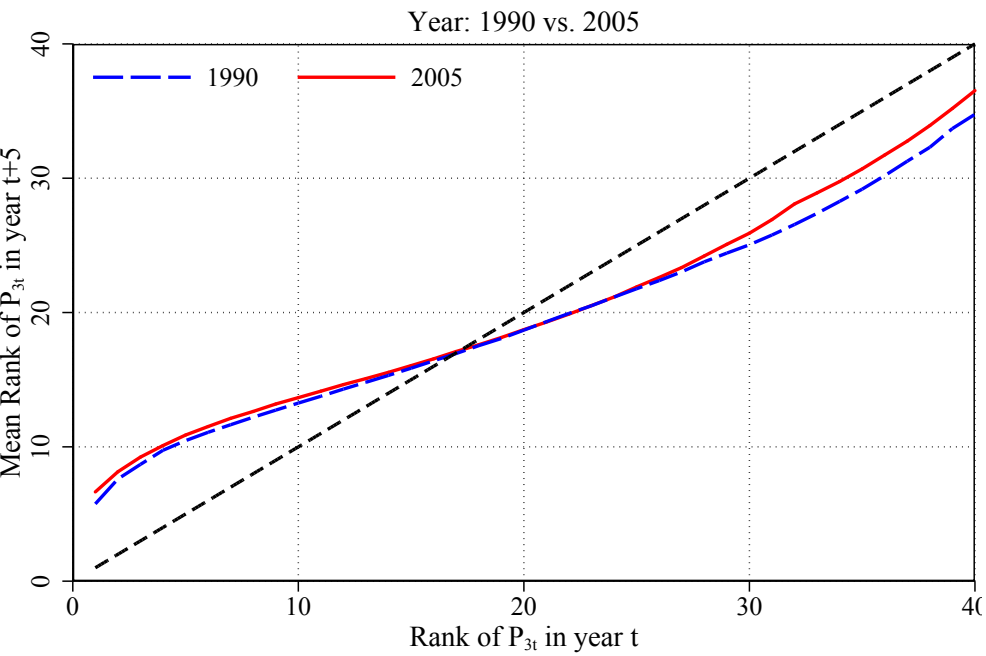


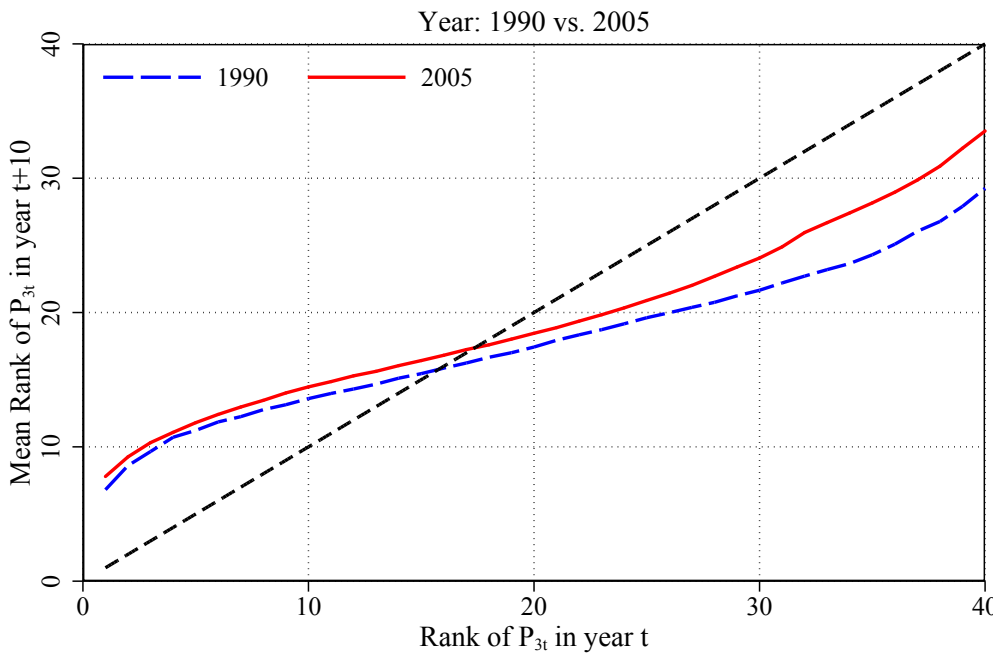


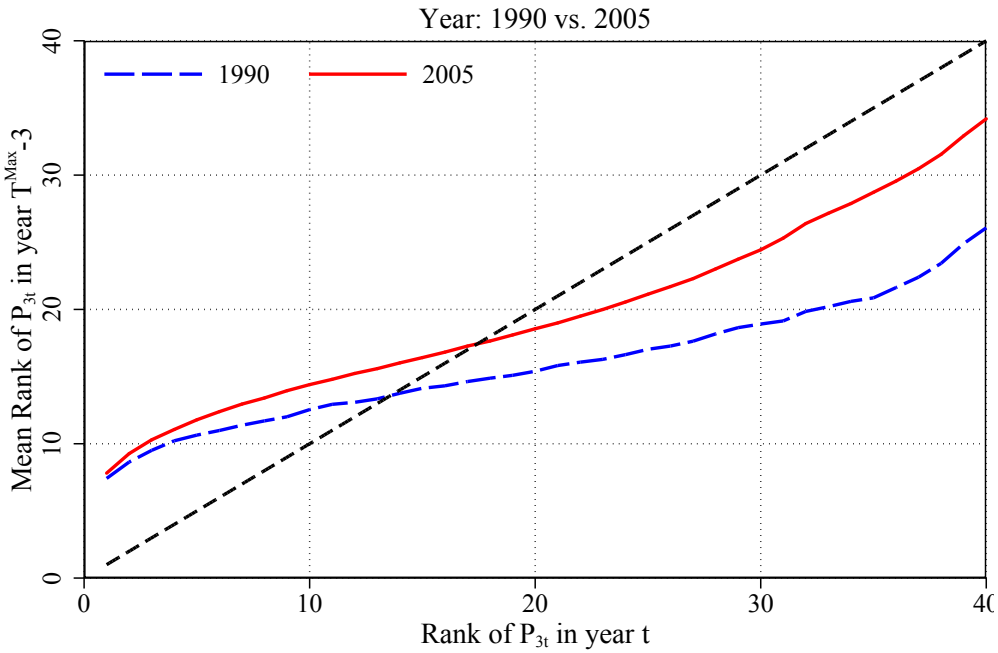


Backup: Perm. Mobility in Residual Inc., 1990 vs. 2005

▸ BACK



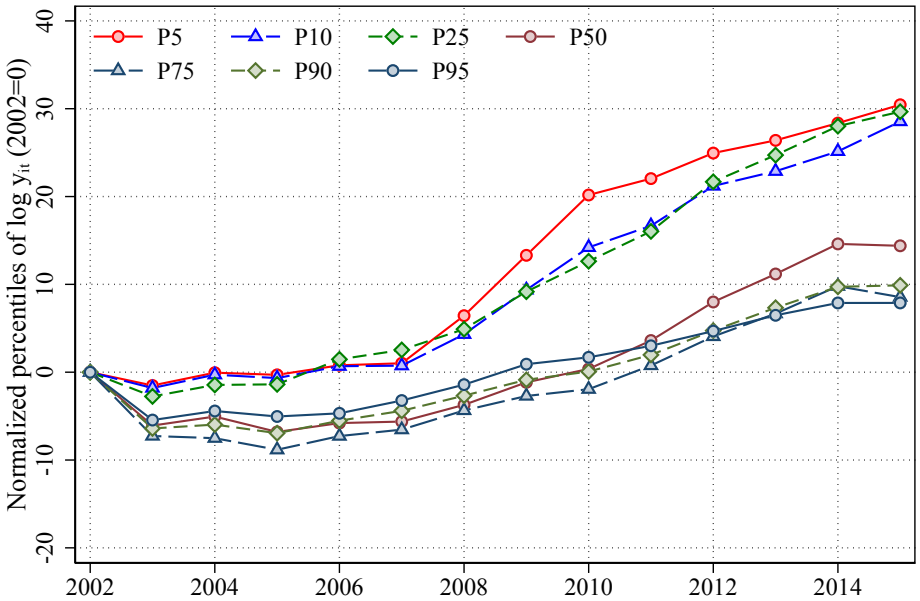




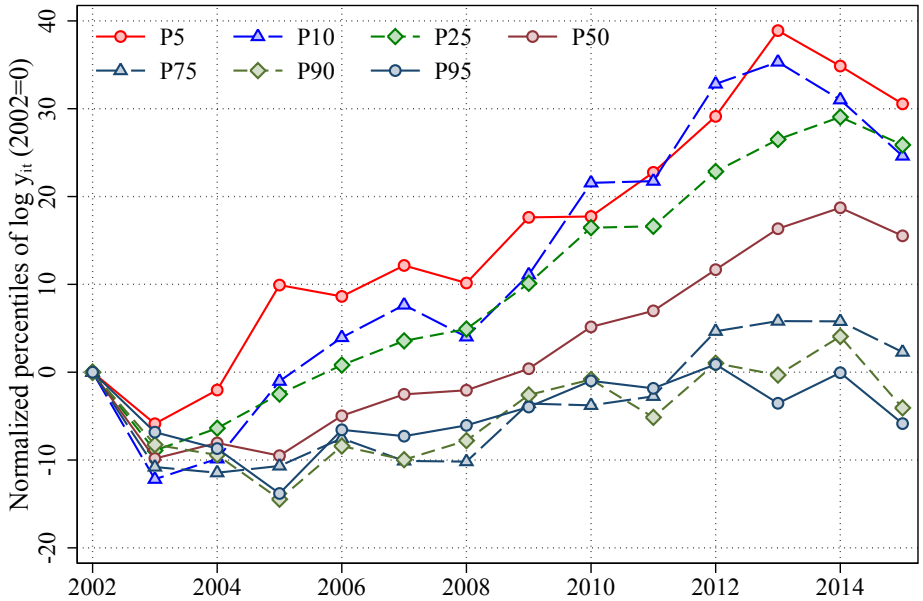
Backup: Norm. Inequ. Statistics in Admin. & Survey Data

▸ Back

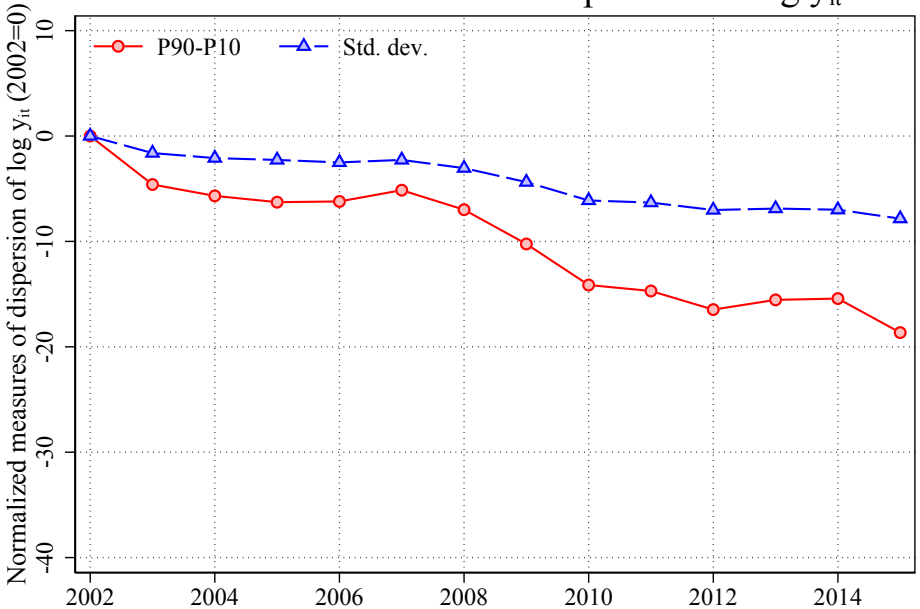
Normalized percentiles of log y_{it}



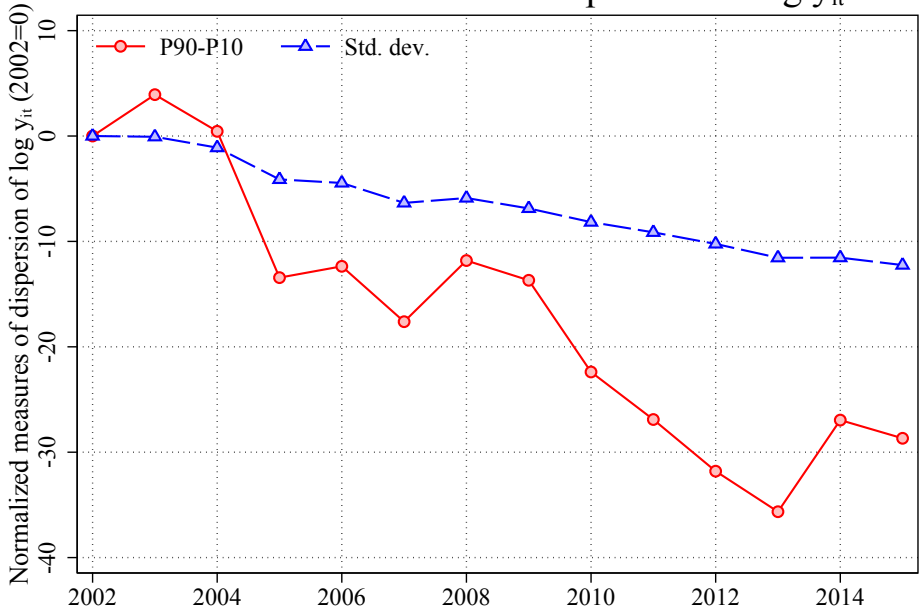
Normalized percentiles of log y_{it}



Normalized measures of dispersion of log y_{it}



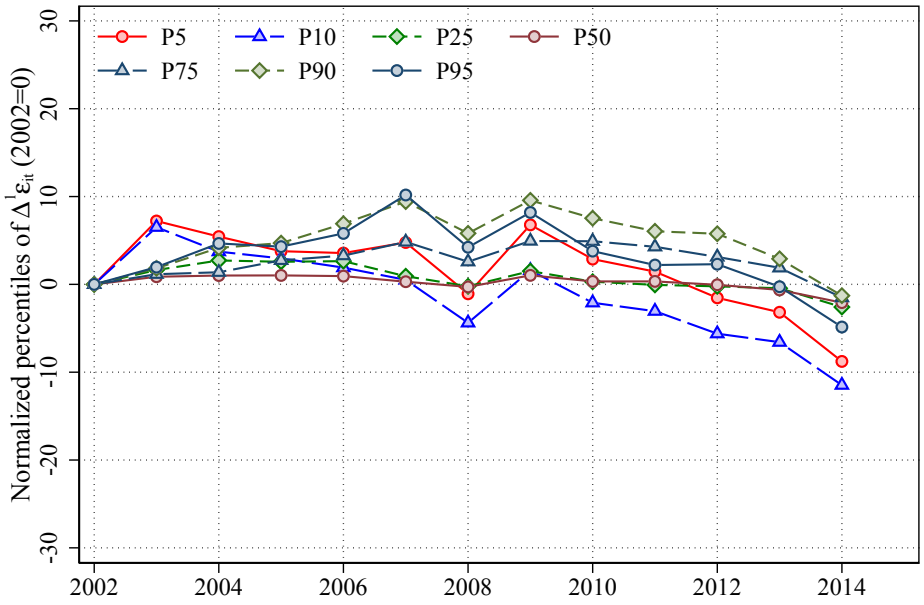
Normalized measures of dispersion of log y_{it}



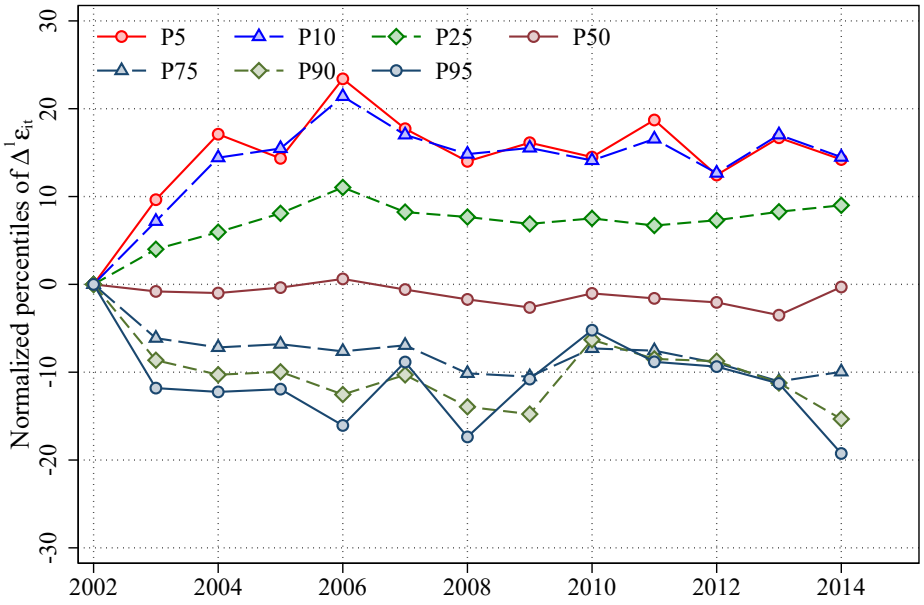
Backup: Norm. Vol. Statistics in Admin. & Survey Data

▸ Back

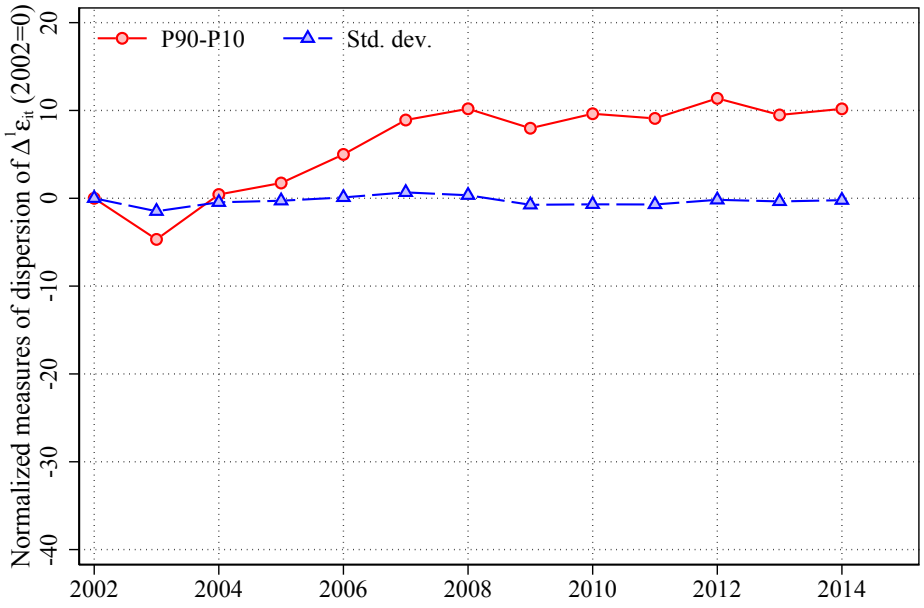
Normalized percentiles of $\Delta^1\varepsilon_{it}$



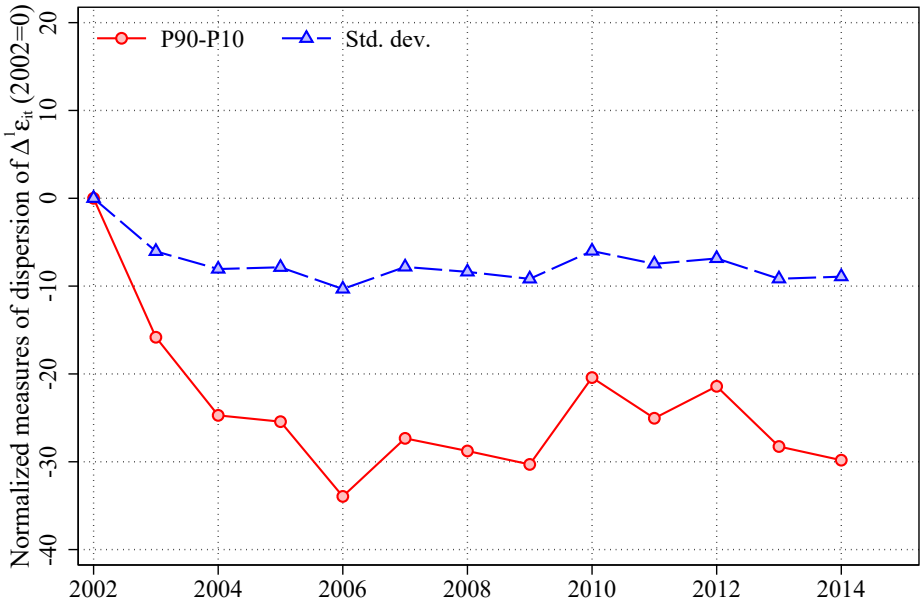
Normalized percentiles of $\Delta^1\varepsilon_{it}$



Normalized measures of dispersion of $\Delta^1\varepsilon_{it}$



Normalized measures of dispersion of $\Delta^1\varepsilon_{it}$

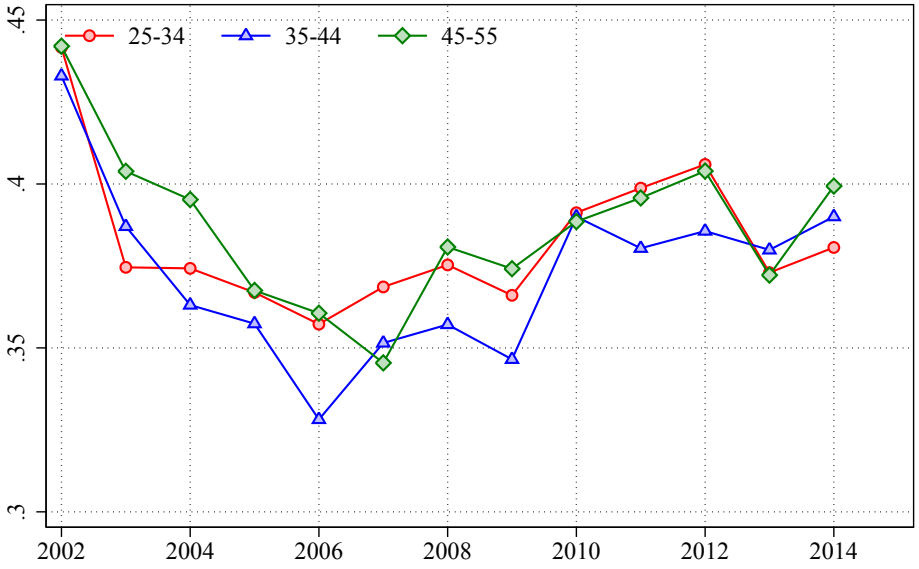


Backup: Sectoral & Within-Sector Changes by Age

▸ Back

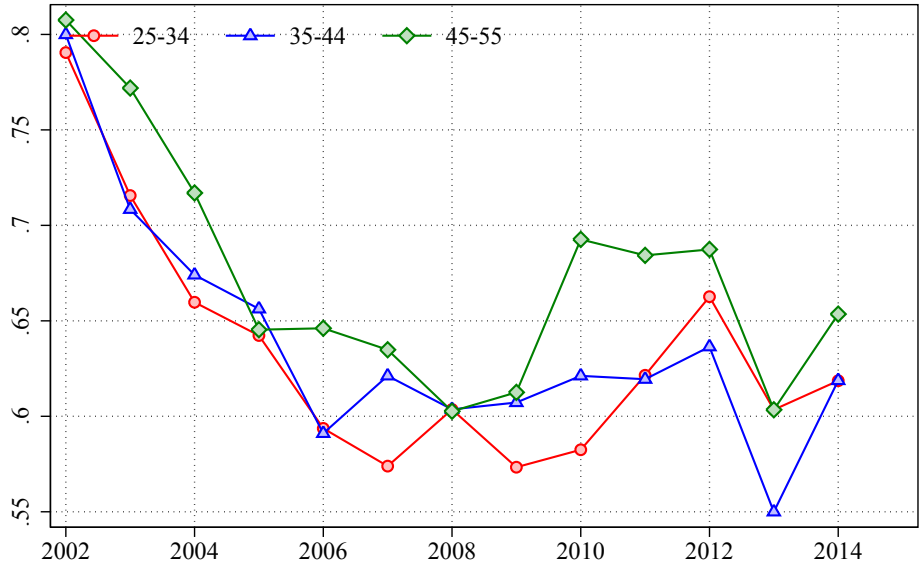
Std of $\Delta^1\varepsilon_{it}$: Formal-Formal

Sample: all, 1-year-forward change



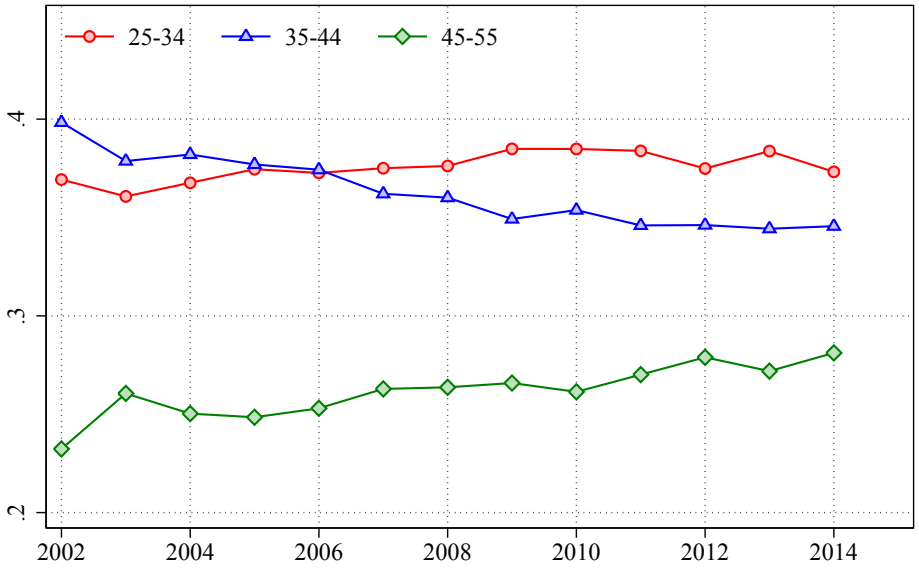
Std of $\Delta^1\varepsilon_{it}$: Informal-Informal

Sample: all, 1-year-forward change



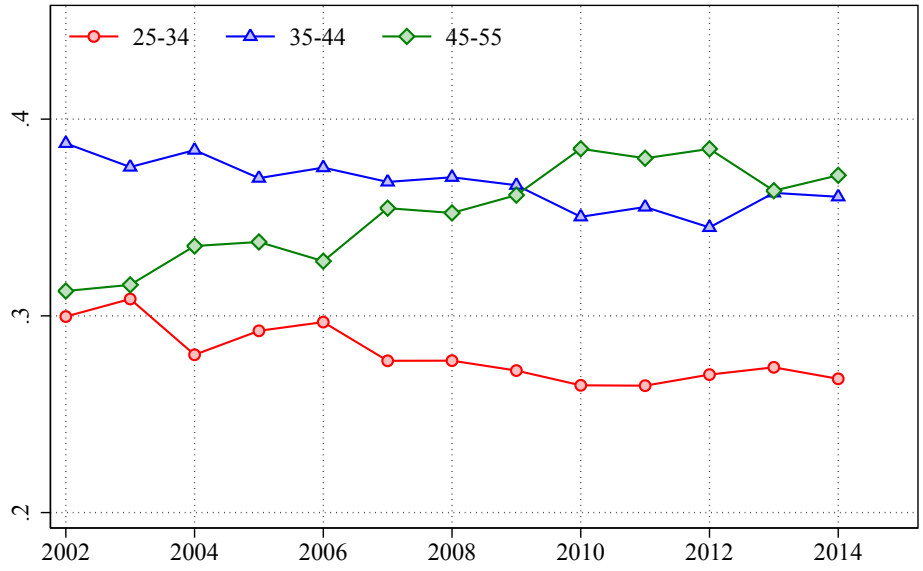
Composition of age group: Formal-Formal

Sample: all, 1-year-forward change



Composition of age group: Informal-Informal

Sample: all, 1-year-forward change

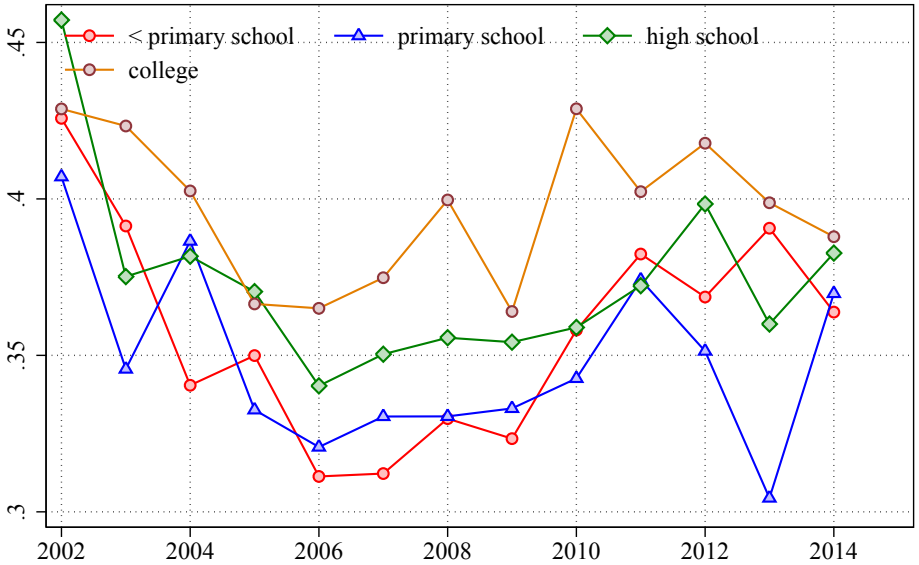


Backup: Sectoral & Within-Sector Changes by Edu

▸ Back

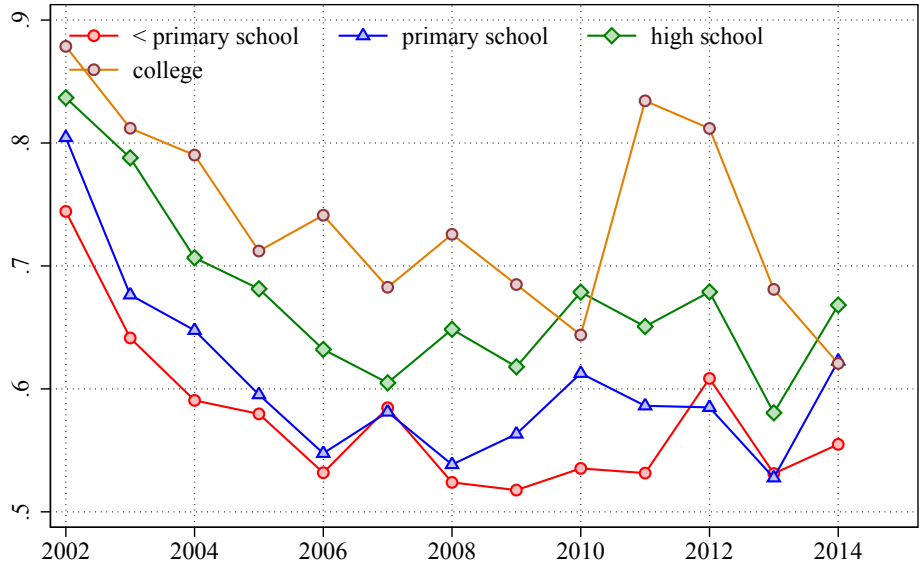
Std of $\Delta^1\varepsilon_{it}$: Formal-Formal

Sample: all, 1-year-forward change



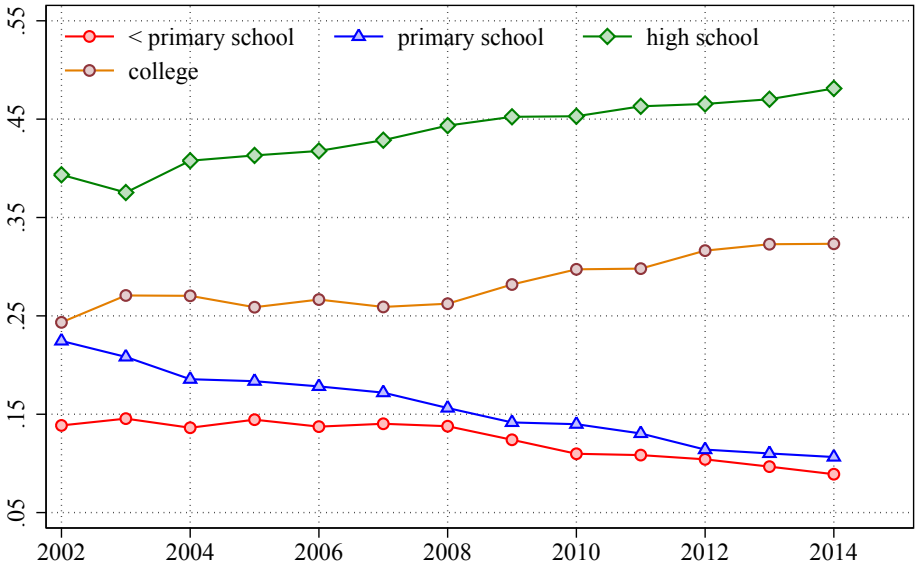
Std of $\Delta^1\varepsilon_{it}$: Informal-Informal

Sample: all, 1-year-forward change



Composition of education degree: Formal-Formal

Sample: all, 1-year-forward change



Composition of education degree: Informal-Informal

Sample: all, 1-year-forward change