How much has wealth concentration grown in the United States? A re-examination of data from 2001-2011

Jesse Bricker, Alice Henriques, and Peter Hansen

Federal Reserve Board

December 15, 2017 (WID - Paris)

Disclaimer

The analysis and conclusions set forth are those of the author and do not indicate concurrence by other members of the research staff or the Board of Governors of the Federal Reserve System.

Presentation Outline

Intro

- How to measure wealth?

Intro

- How to measure wealth?
- Household surveys
  - Survey of Consumer Finances (SCF)

Intro

- How to measure wealth?
- Household surveys
  - Survey of Consumer Finances (SCF)
    - Fairly recent (SCF 1983-; EFF 2001-; HFCS 2010-)
    - Oversample for credible representation at top
    - $\Longrightarrow$ A mix of admin. and survey data

Intro

- How to measure wealth?
- Household surveys
  - Survey of Consumer Finances (SCF)
    - Fairly recent (SCF 1983-; EFF 2001-; HFCS 2010-)
    - Oversample for credible representation at top
    - $\Longrightarrow$ A mix of admin. and survey data
- Administrative data
  - No U.S. administrative *wealth* data, *impute* from income tax
    - Saez + Zucman, 2016, Greenwood, 1983

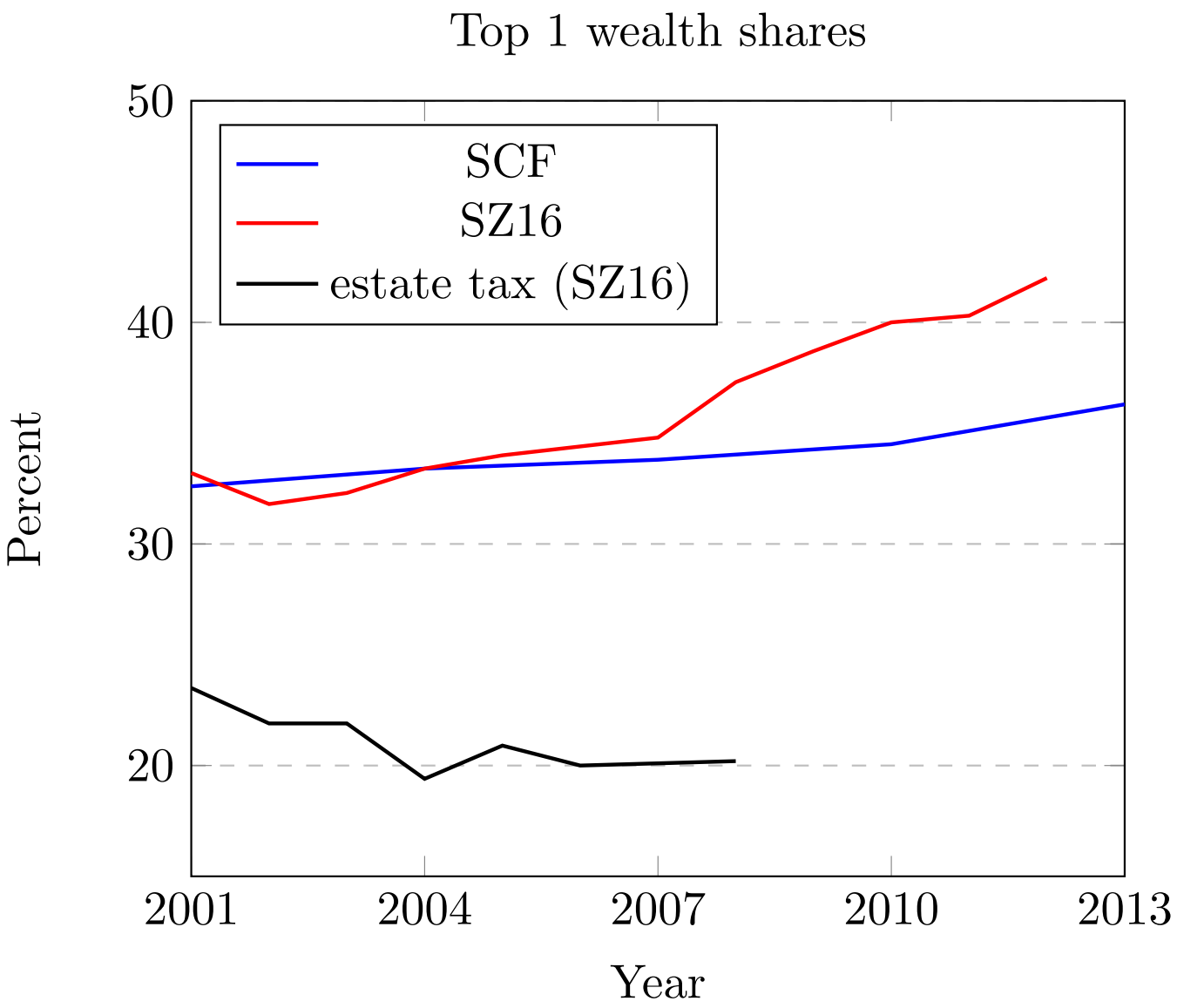
Intro

- How to measure wealth?
- Household surveys
  - Survey of Consumer Finances (SCF)
    - Fairly recent (SCF 1983-; EFF 2001-; HFCS 2010-)
    - Oversample for credible representation at top
    - $\Longrightarrow$ A mix of admin. and survey data
- Administrative data
  - No U.S. administrative *wealth* data, *impute* from income tax
    - Saez + Zucman, 2016, Greenwood, 1983
    - Income tax data: excellent coverage at top of income distribution
    - Piketty 1999; Piketty + Saez, 2003; Alvaredo + Saez, 2009

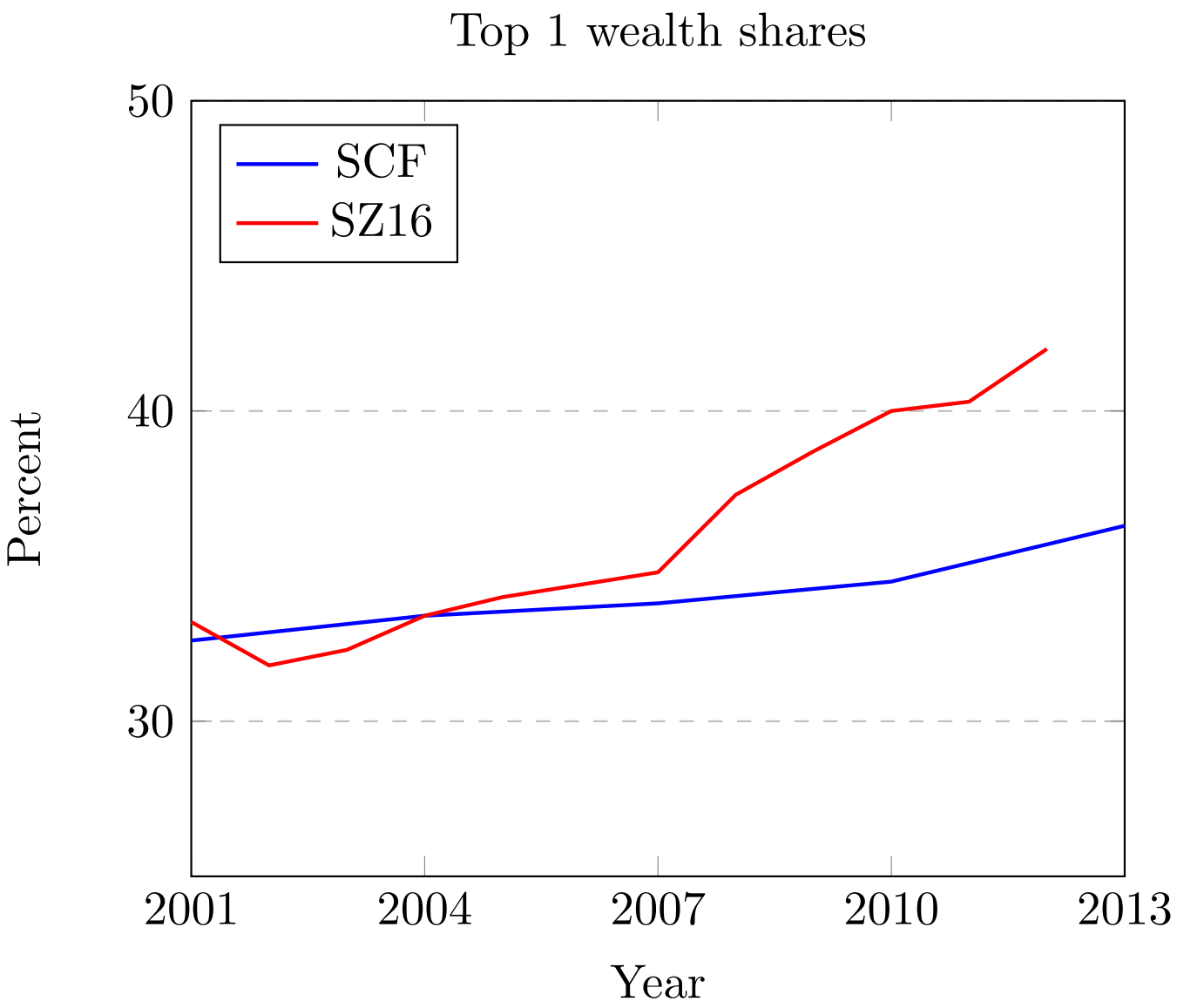
Intro

- How to measure wealth?
- Household surveys
  - Survey of Consumer Finances (SCF)
    - Fairly recent (SCF 1983-; EFF 2001-; HFCS 2010-)
    - Oversample for credible representation at top
    - $\Longrightarrow$ A mix of admin. and survey data
- Administrative data
  - No U.S. administrative *wealth* data, *impute* from income tax
    - Saez + Zucman, 2016, Greenwood, 1983
    - Income tax data: excellent coverage at top of income distribution
    - Piketty 1999; Piketty + Saez, 2003; Alvaredo + Saez, 2009
  - Estate tax filings (Kopczuk + Saez, 2003, only very top, no updates)

Different growth rates: SCF vs. SZ16 vs. estate tax

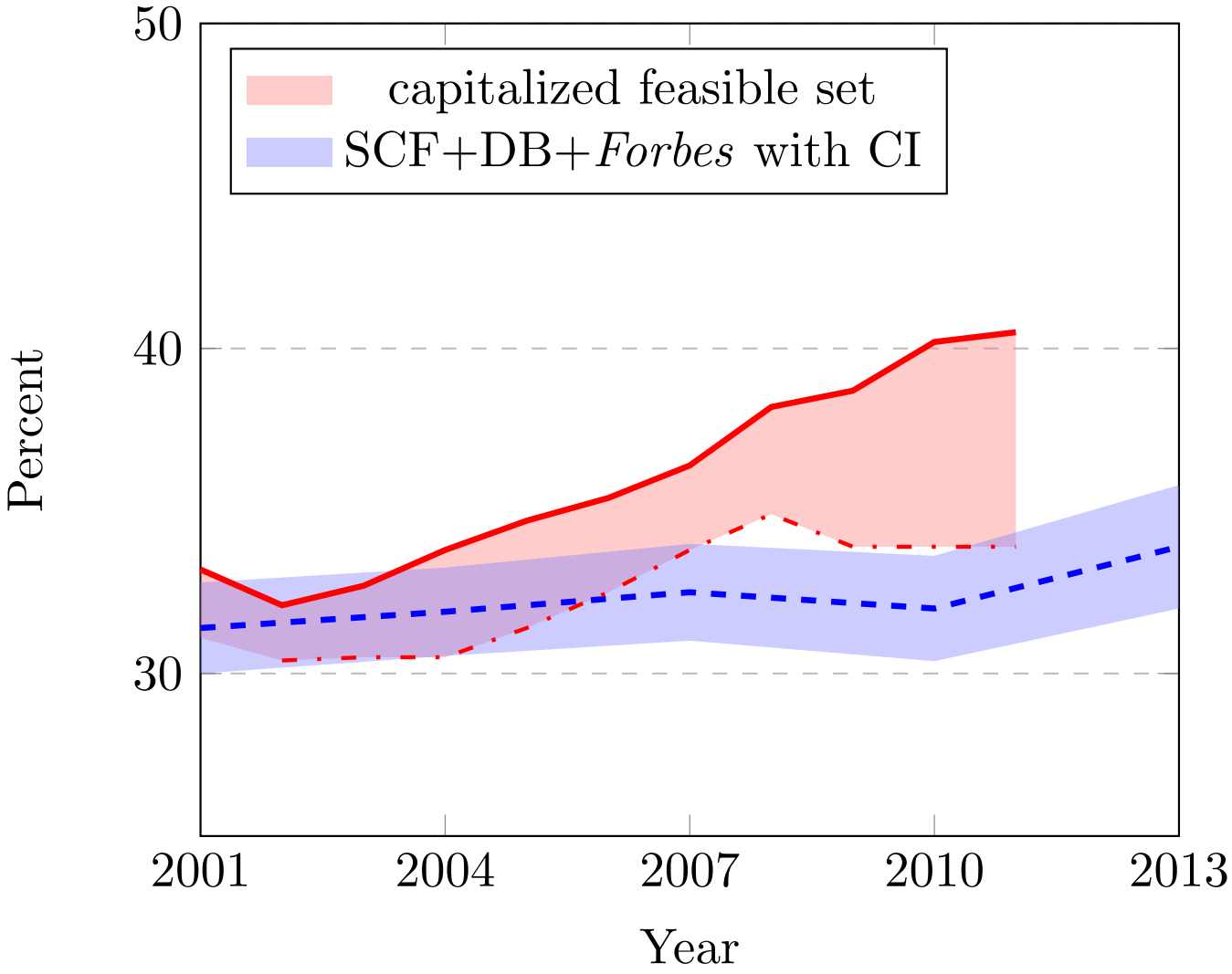


Different growth rates: SCF vs. SZ16



Large variability in growth imputed wealth concentration

SCF and capitalized top 1 wealth shares with uncertainty

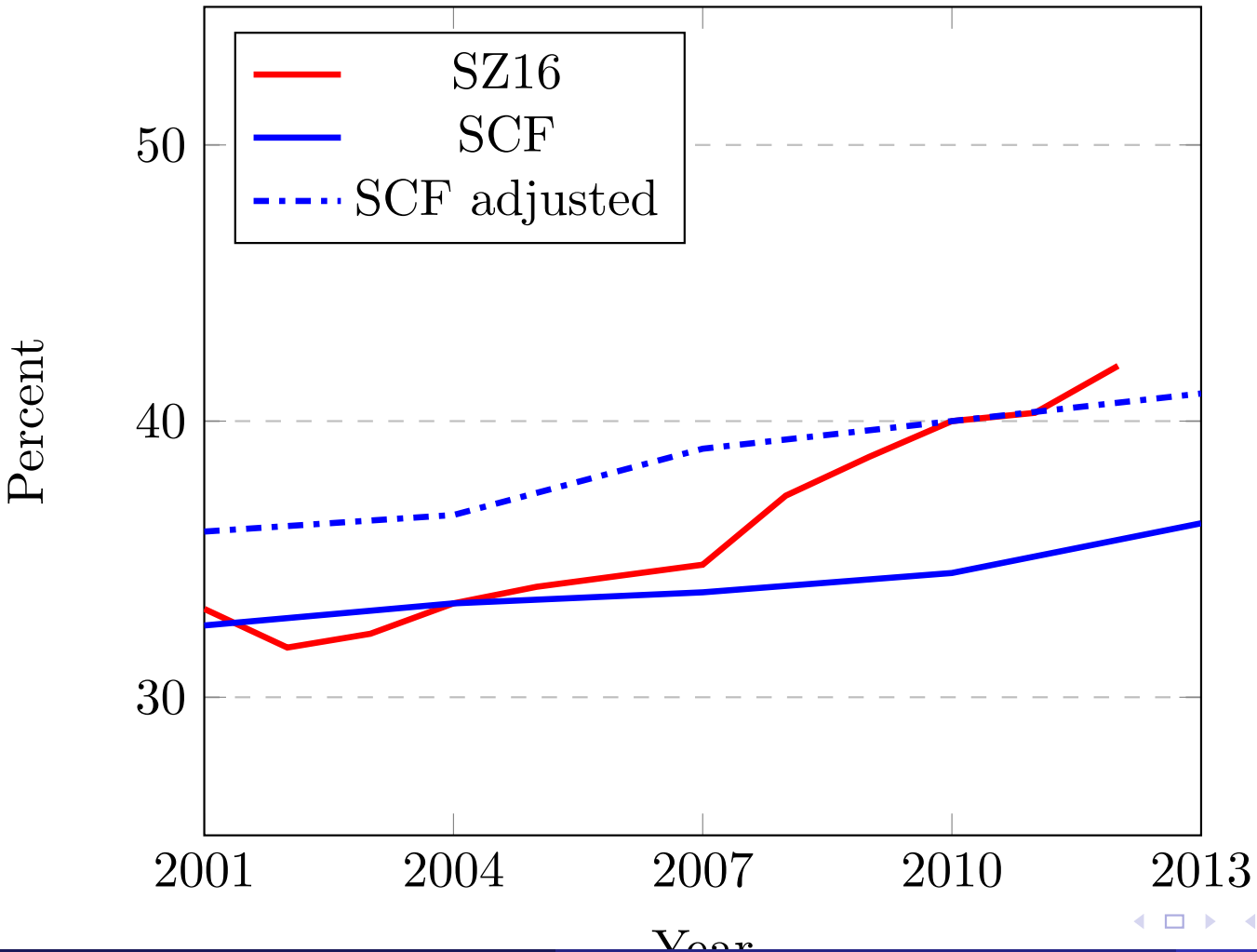


Presentation Outline

1. Intro
2. Can measurement differences explain differences in growth?
3. Variability
   - SCF - household survey
   - Income tax data - infer wealth
   - Sensitivity
   - SCF survey
   - Aside - alternate survey coverage error proxy

Measurement differences cannot explain trend differences

- unit of observation (tax unit vs. family)
- how measured (imputed vs. self-reported)
- concepts (DB vs. no DB)



Measurement

- This is covered in detail Bricker et al (2016, BPEA)
- Here:
  - Reconcile growth differences with *sensitivity* of estimates?

Presentation Outline

1. Intro
2. Can measurement differences explain differences in growth?
3. Variability
   - SCF - household survey
   - Income tax data - infer wealth
   - Sensitivity
   - SCF survey
   - Aside - alternate survey coverage error proxy

(1) SCF survey

- A household survey with *wealthy oversample*
  - Total sample size $\approx$ 6,500, incl. wealthy *oversample* $\approx$ 1,500 families
- Wealthy oversample
  - Based on a *sample* of admin. records derived from income tax returns
  - Use two models: rank order families by expected wealth

(1) SCF survey

- A household survey with *wealthy oversample*
  - Total sample size $\approx$ 6,500, incl. wealthy *oversample* $\approx$ 1,500 families
- Wealthy oversample
  - Based on a *sample* of admin. records derived from income tax returns
  - Use two models: rank order families by expected wealth
- **Model 1**: capitalize income (inflate by rate of return)
  - $\hat{wealth}_i = \hat{nonfin}_i + kg_i + \sum_{k=1}^{K}(income_i^k / return^k), k = 1...7$

(1) SCF survey

- A household survey with *wealthy oversample*
  - Total sample size $\approx$ 6,500, incl. wealthy *oversample* $\approx$ 1,500 families
- Wealthy oversample
  - Based on a *sample* of admin. records derived from income tax returns
  - Use two models: rank order families by expected wealth
- **Model 1**: capitalize income (inflate by rate of return)
  - $\hat{wealth}_i = \hat{nonfin}_i + kg_i + \sum_{k=1}^{K}(income_i^k / return^k), k = 1...7$
- **Model 2**: correlate income and wealth
  - Use capital, wage, pension, etc... income

(1) SCF survey

- A household survey with *wealthy oversample*
  - Total sample size $\approx$ 6,500, incl. wealthy *oversample* $\approx$ 1,500 families
- Wealthy oversample
  - Based on a *sample* of admin. records derived from income tax returns
  - Use two models: rank order families by expected wealth
- **Model 1**: capitalize income (inflate by rate of return)
  - $\hat{wealth}_i = \hat{nonfin}_i + kg_i + \sum_{k=1}^{K}(income_i^k / return^k), k = 1...7$
- **Model 2**: correlate income and wealth
  - Use capital, wage, pension, etc... income
- Select sample of $\approx$ 5,100 ($\approx$ 1,500 respond)
  - Majority are in top 1 pct.
  - Easily identifiable thrown out (e.g. *Forbes* 400)

(2) Income tax data - infer wealth

- **Use same data as SCF oversample**
  - A *sample* of administrative records derived from income tax returns
  - E.g. Saez Zucman (2016) Greenwood (1983)
  - Here: the PUF with Saez (2016) supplement to match INSOLE

(2) Income tax data - infer wealth

- **Use same data as SCF oversample**
  - A *sample* of administrative records derived from income tax returns
  - E.g. Saez Zucman (2016) Greenwood (1983)
  - Here: the PUF with Saez (2016) supplement to match INSOLE
- Use *one* (capitalization) model to rank families and predict wealth
  - $\hat{wealth}_i = \hat{nonfin}_i + kg_i + \sum_{k=1}^{K}(income_i^k / return^k), k = 1...7$
  - Align to known distributions of wages, pensions; use *some* debts

(2) Income tax data - infer wealth

- **Use same data as SCF oversample**
  - A *sample* of administrative records derived from income tax returns
  - E.g. Saez Zucman (2016) Greenwood (1983)
  - Here: the PUF with Saez (2016) supplement to match INSOLE
- Use *one* (capitalization) model to rank families and predict wealth
  - $\hat{wealth}_i = \hat{nonfin}_i + kg_i + \sum_{k=1}^{K}(income_i^k / return^k), k = 1...7$
  - Align to known distributions of wages, pensions; use *some* debts
- What **rate of return**?
  - Ratio: taxed income flow to FA asset stock? (Saez + Zucman, 2016)
  - Market rates? (Greenwood, 1983, Bricker, Henriques, Moore, 2017)
  - Heterogeneous returns? (Fagereng et al 2016)

What can go wrong?

	Modeled income tax data	Survey (e.g. SCF)
Coverage error	Yes	Yes
Sampling error	Yes	Yes
Unit nonresp error	Yes	Yes
Item nonresp error	Yes	Yes
Adjustment error	Yes	Yes
Concept validity	Yes	Yes
Measurement error	Yes	Yes
Processing error	Yes	Yes
Model error	Yes	No

Table: Potential errors

Model uncertainty - income tax data

- Heterogeneous rate of return on fixed income for top 1%
  - Uncertainty of estimates is partially discussed in SZ16
  - Interest income RoRs become very small in late 2000s
  - $\Longrightarrow$ Almost all growth in concentration due to fixed income assets

Model uncertainty - income tax data

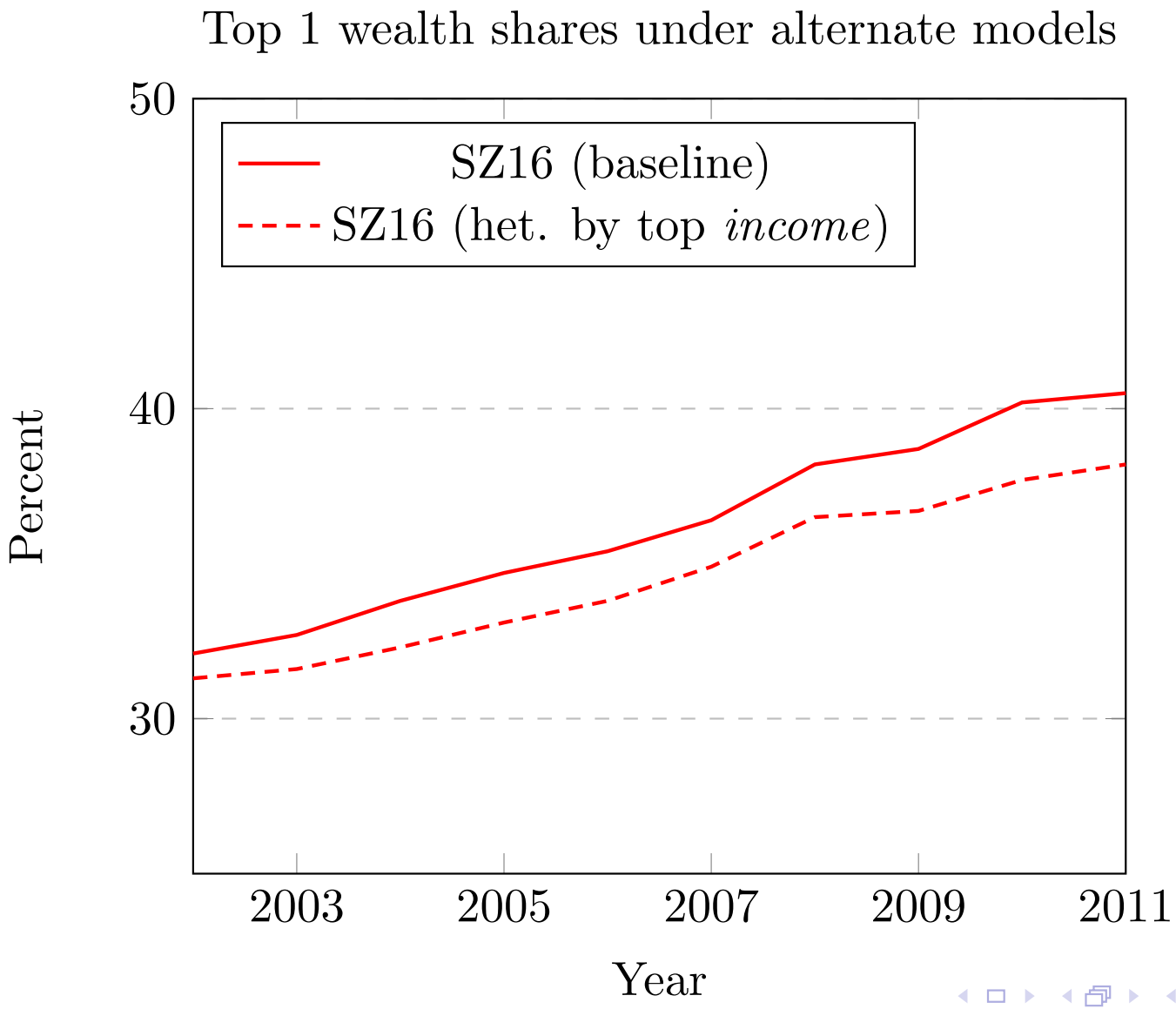
- Heterogeneous rate of return on fixed income for top 1%
  - Uncertainty of estimates is partially discussed in SZ16
  - Interest income RoRs become very small in late 2000s
  - $\Longrightarrow$ Almost all growth in concentration due to fixed income assets
- How?
  - Let top 1 have RoR on fixed income of 10-year Treasury
- Which top 1? Total income, interest income, wealth...?
- It matters!

Model uncertainty - income tax data

- Heterogeneous rate of return on fixed income for top 1%
  - Uncertainty of estimates is partially discussed in SZ16
  - Interest income RoRs become very small in late 2000s
  - $\Longrightarrow$ Almost all growth in concentration due to fixed income assets
- How?
  - Let top 1 have RoR on fixed income of 10-year Treasury
- Which top 1? Total income, interest income, wealth...?
- It matters!
- *Note: just focusing one one type of asset (fixed income)!*

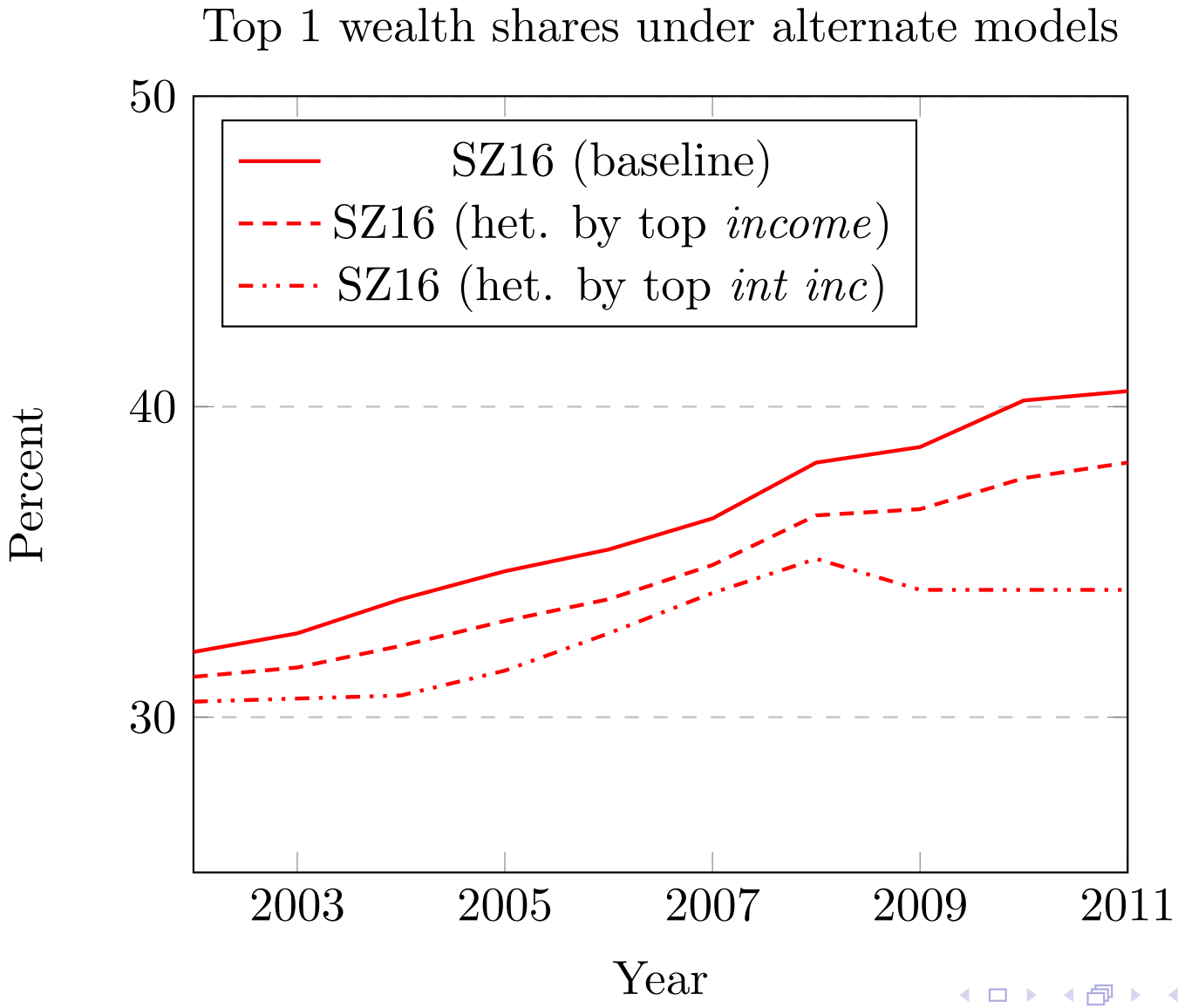
Model uncertainty - income tax data

- Heterogeneous return for top 1 of *total income*



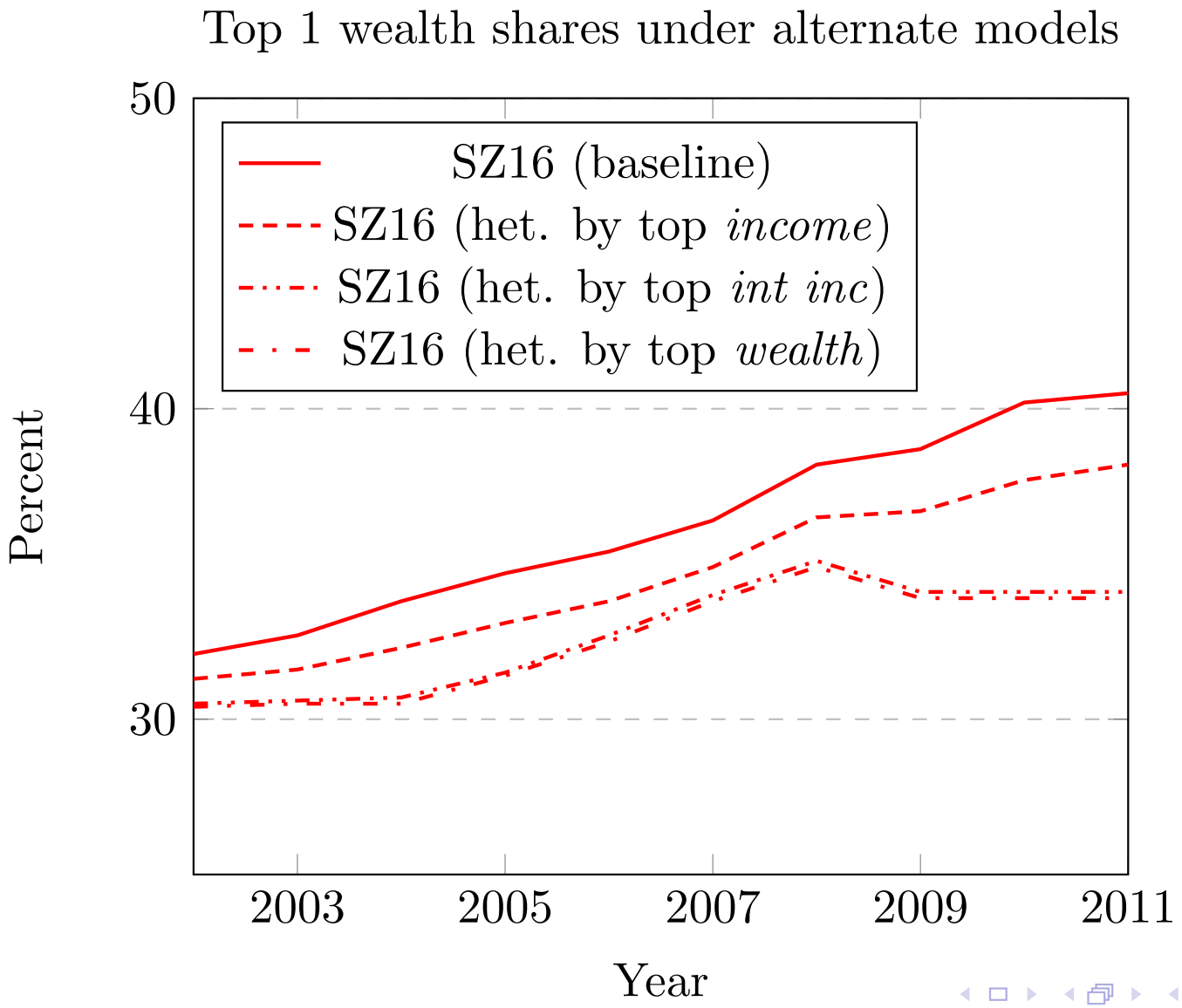
Model uncertainty - income tax data

- Heterogeneous return for top 1 of *interest income*



Model uncertainty - income tax data

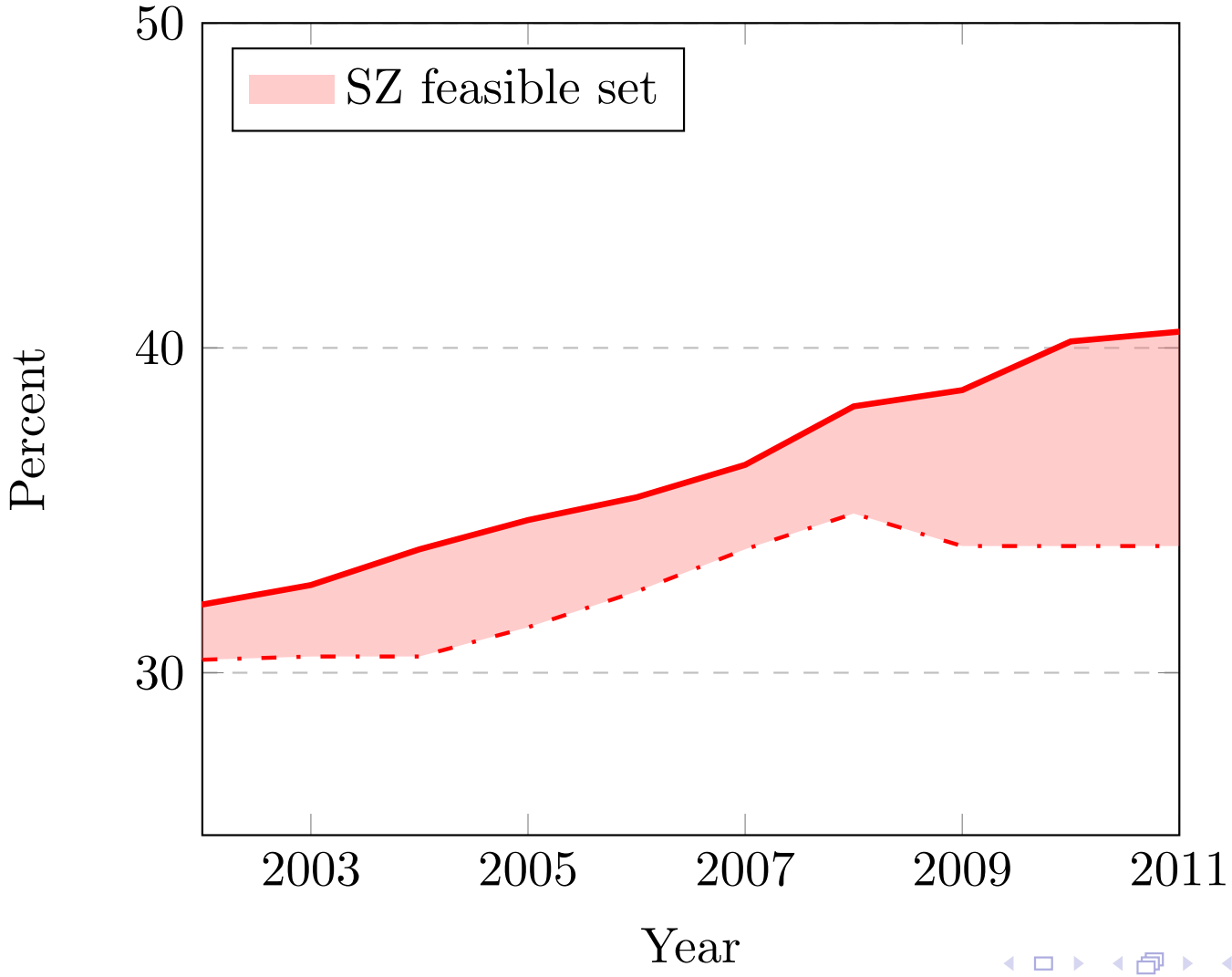
- Heterogeneous return for top 1 of *total wealth* (iterative...)



Model uncertainty - income tax data

- Feasible region: heterogeneous retun on *fixed income assets only*

Top 1 wealth shares under alternate models of fixed income



Evidence for heterogeneous rates of return

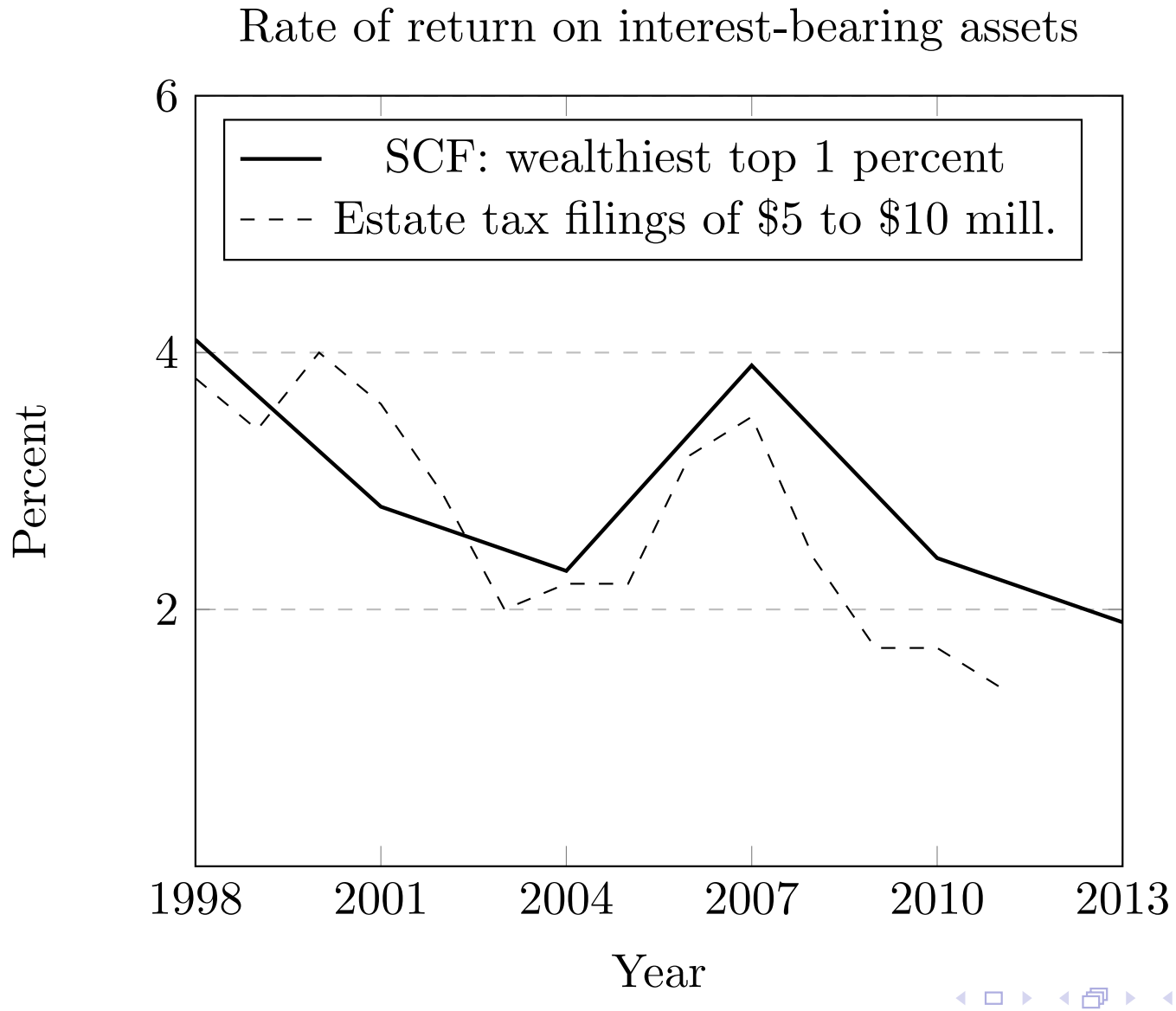
- Heterogeneous returns matter, especially by wealth.
  - Baseline growth: 31.8 to 40.3 (8.5 ppts.)
  - Heter. return *wealth*: 30.3 to 33.9 (3.6 ppts)
- What is the evidence that they exist?
  - Norwegian registry data: Fagereng et al, (2016)
  - United States data: below

	Bottom 99	Top 1	Ratio (top:bot)
SZ16	1.2%	≈2%	1.67
SCF	1.4%	2.4%	1.71
SCF (Model 2)	3.0%	5.4%	1.80

Table: Rates of return on interest-bearing assets (2010)

Evidence for heterogeneous rates of return

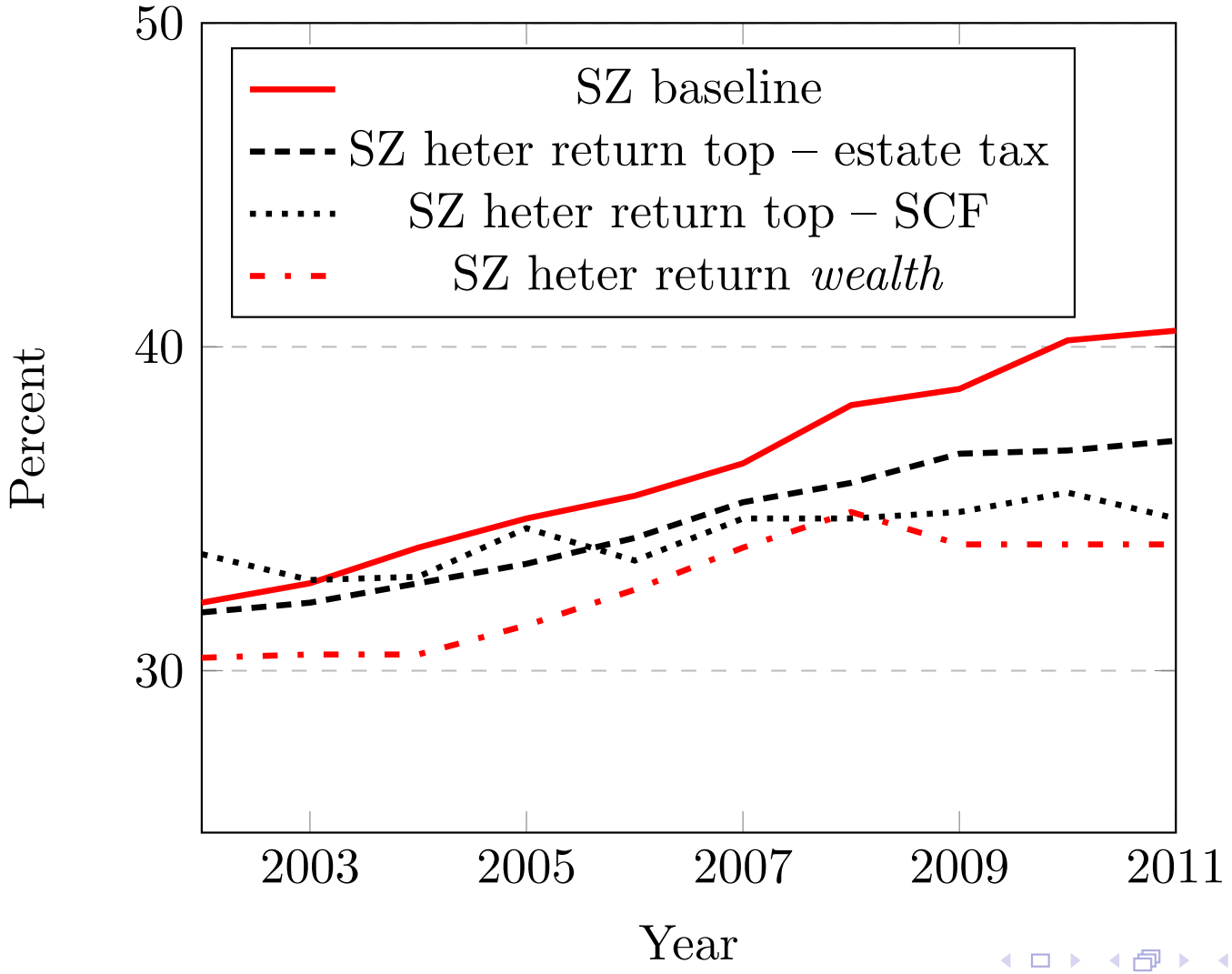
- The same time trend



Model uncertainty - alternate heterogeneous returns

- Use estate tax-income tax, SCF top rate of return

Top 1 wealth shares: estate tax, SCF interest rates of return



Now for survey...

	Survey (e.g. SCF)	Income tax data
Coverage error	Yes	Yes
Sampling error	Yes	Yes
Unit nonresp error	Yes	Yes
Item nonresp error	Yes	Yes
Adjustment error	Yes	Yes
Concept validity	Yes	Yes
Measurement error	Yes	Yes
Processing error	Yes	Yes
Model error	No	Yes

Table: Potential errors

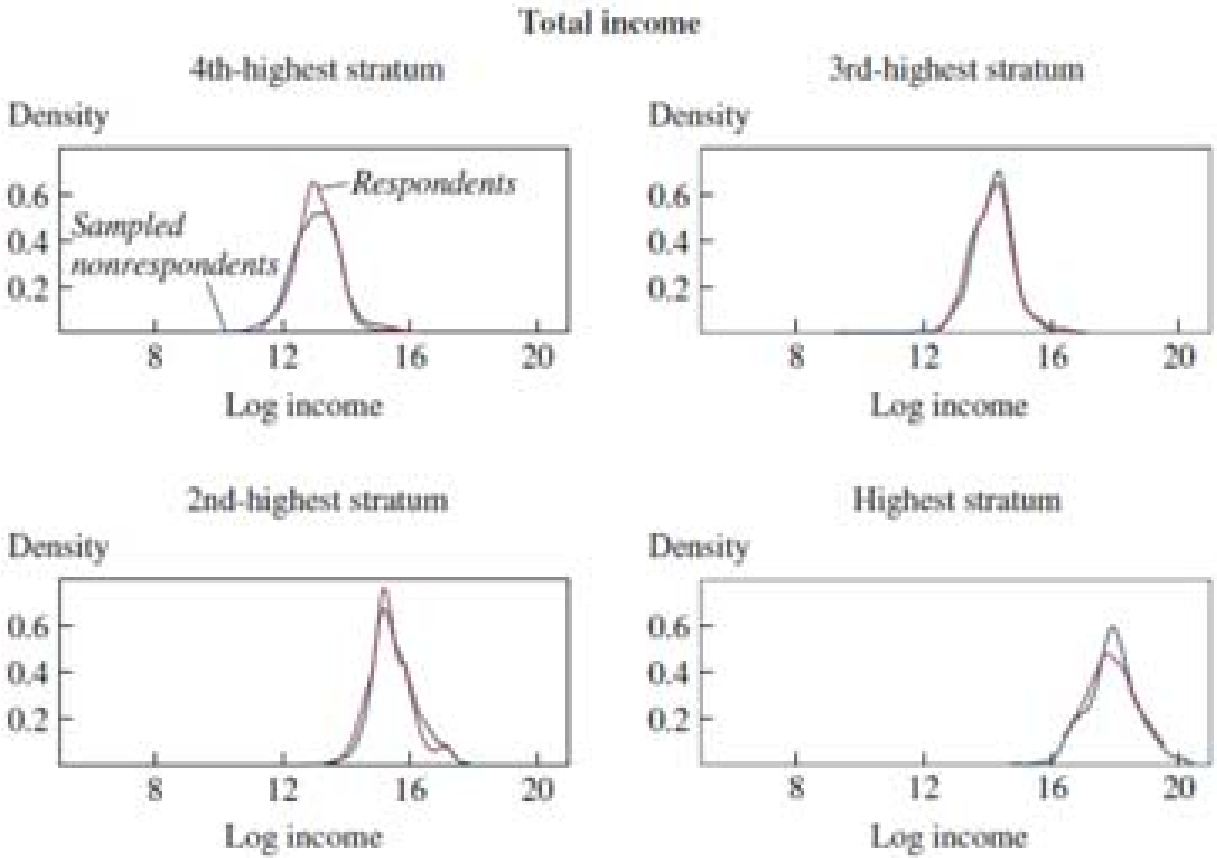
Can focus on these four (get LB!)

	Survey (e.g. SCF)	Income tax data
Coverage error	Yes	Yes
Sampling error	Yes	Yes
Unit nonresp error	Yes	Yes
Item nonresp error	Yes	Yes
Adjustment error	Yes	Yes
Concept validity	Yes	Yes
Measurement error	Yes	Yes
Processing error	Yes	Yes
Model error	No	Yes

Table: Potential errors

Survey sensitivity : unit nonresponse error

- Unit nonresponse error? Ignore for now...



Survey sensitivity : coverage error

- SCF precluded from sampling *Forbes* 400
  - But good evidence that SCF gets whole distribution, barring *Forbes* (Bricker et al 2016)

Survey sensitivity : coverage error

- SCF precluded from sampling *Forbes* 400
  - But good evidence that SCF gets whole distribution, barring *Forbes* (Bricker et al 2016)
- Add *Forbes* 400 wealth by modifying sample weights
  - Overlap between SCF respondents, *Forbes* (e.g. Vermeulen 2015)
  - How? because non-public wealth, imperfect *Forbes*

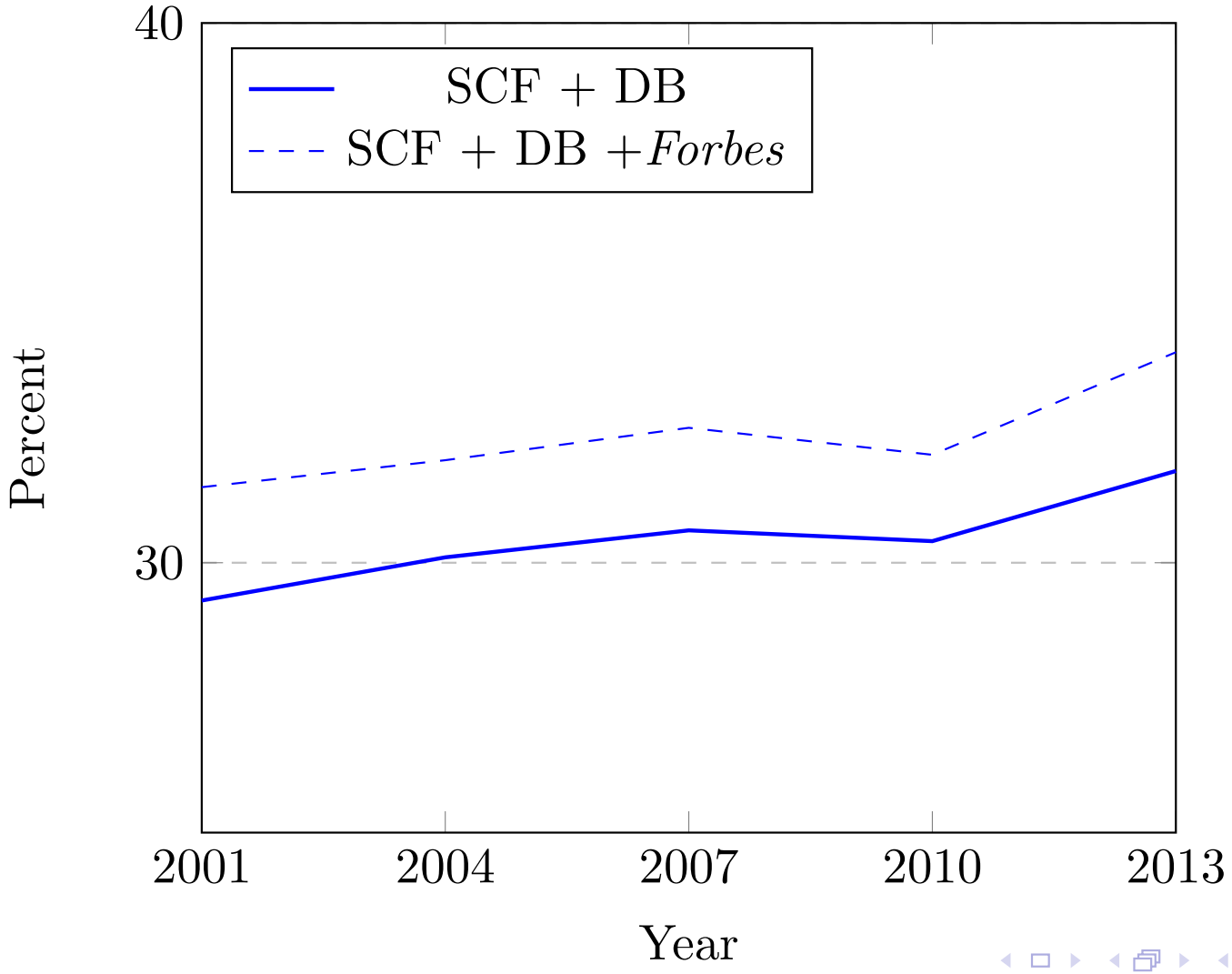
Survey sensitivity : coverage error

- SCF precluded from sampling *Forbes* 400
  - But good evidence that SCF gets whole distribution, barring *Forbes* (Bricker et al 2016)
- Add *Forbes* 400 wealth by modifying sample weights
  - Overlap between SCF respondents, *Forbes* (e.g. Vermeulen 2015)
  - How? because non-public wealth, imperfect *Forbes*
- *Combine samples* using overlap between Forbes and SCF respondents
  - Each Forbes unit is one self-representing family (weight=1)
  - Bin net worth. Adjust Forbes, SCF weights by relative frequency
  - Similar to List sample + AP are blended in SCF weighting

Survey sensitivity : coverage error

- Mostly a level shift up

SCF top shares, including and excluding *Forbes* wealth



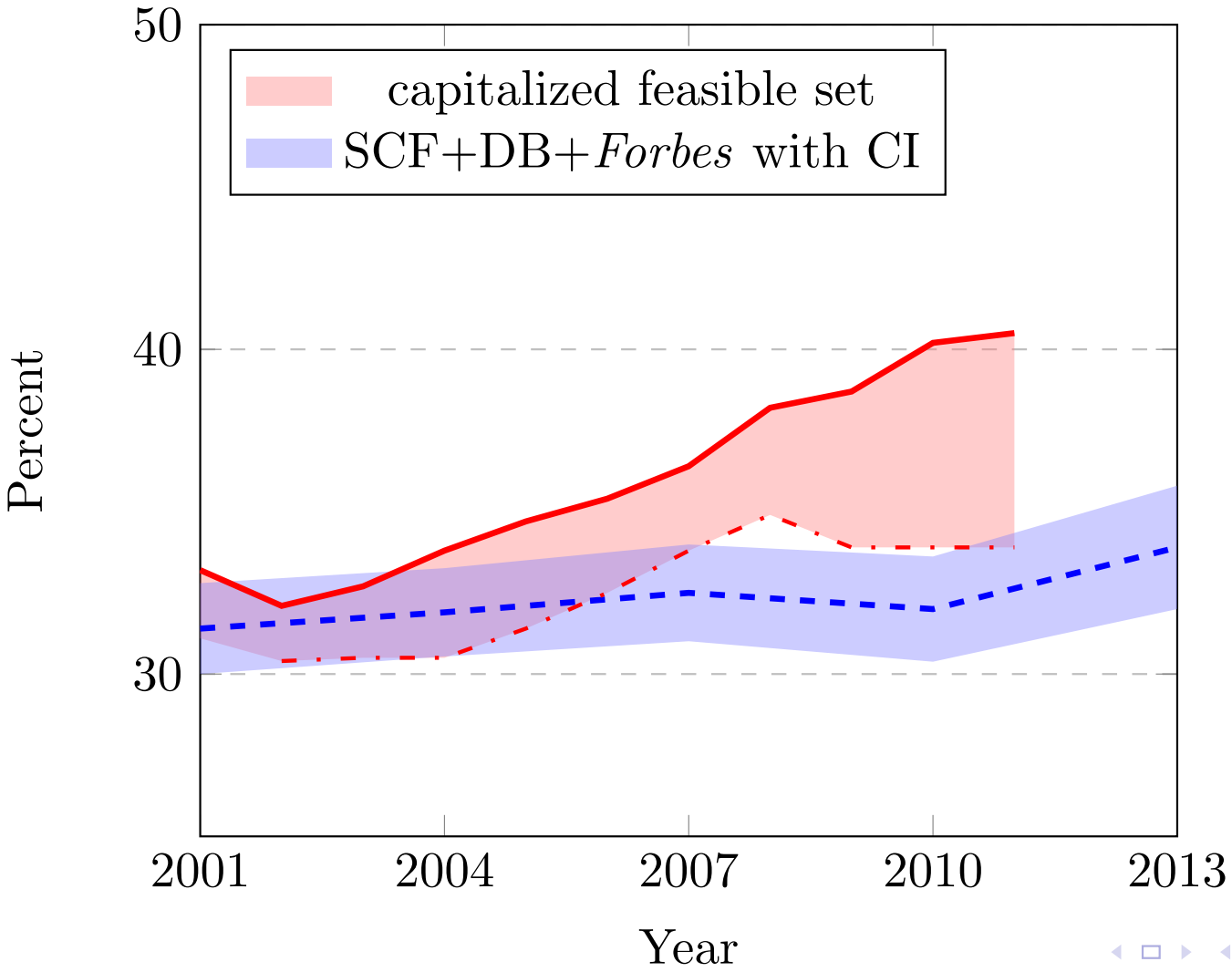
Survey sensitivity : sampling, item nonresponse

- Sampling error
  - Use bootstrap replication techniques as proxy for sampling variablity
- Item nonresponse error
  - Imputation variance as proxy for variablity due to item nonresponse

Sensitivity of survey, capitalized wealth

- Control for sampling error, item nonresponse error, coverage error

SCF and capitalized top 1 wealth shares with uncertainty



Concluding

- Sensitivity in wealth concentration estimates
- SCF survey: estimate *(some)* sampling and non-sampling variability
  - Introduce weights to blend *Forbes*, reduce coverage error
- Imputed wealth from income: demonstrate modeling variability
  - Potentially very large, negating the benefits of good top end coverage
  - Did not estimate sampling and non-sampling variability (next step)
- *No difference* in growth of wealth concentration...
  - ...if believe that top 1 of *total wealth* or *interest income* should get larger interest rate of return (not top 1 of *total income*)

Extra slides

Pareto: an alternate survey coverage error proxy

- Estimate model parameters via ML

$$\hat{\alpha_{ml}} = [\sum_{w_i=w_{min}}^{w_{max}} n(w_i)/N(w_{min}) * ln(w_i/w_{min})]^{-1},$$

$$SE = \hat{\alpha_{ml}} * 1/(N(w_{min})^{-0.5})$$

- Estimate of $\hat{\alpha}$ depends on where x_{min} begins
- Where does x_{min} begin?

Rich list + SCF

x_{min}	α	SCF + Rich list Top1 share
\$4m	1.51	37.7
\$5m	1.53	37.3
\$10m	1.74	35.0
\$15m	1.68	34.9

Table: 2010 SCF + Rich list

- Replace SCF data above these cutoffs with Pareto interpolation, re-estimate wealth shares

Rich list + SCF

- Assume x_{min} begins at \$4m, \$5m, \$10m, \$15m
- Replace SCF data above these cutoffs with Pareto interpolation, re-estimate wealth shares

Rich list + SCF sensitivity (2)