Using Income to Predict Wealth: an Exploration with Panel Data

Jesse Bricker and Alice Henriques

Federal Reserve Board

July 19, 2016

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

1. Intro
2. Infer wealth
3. Noisy identification of the top?
4. Estimate persistence at the top?
5. Outside the top

Intro

- Three main ways to measuring wealth

Intro

- Three main ways to measuring wealth
- Surveys
  - Survey of Consumer Finances (SCF), 1989-current
  - Good national and top representation, via oversample
  - e.g. Bricker, Henriques, Krimmel, Sabelhaus (2016)

Intro

- Three main ways to measuring wealth
- Surveys
  - Survey of Consumer Finances (SCF), 1989-current
  - Good national and top representation, via oversample
  - e.g. Bricker, Henriques, Krimmel, Sabelhaus (2016)
- Estate tax data
  - Infer wealth from tax filing by modeling mortality
  - e.g. Kopczuk and Saez (2004), Johnson and Raub (2015)

Intro

- Three main ways to measuring wealth
- Surveys
  - Survey of Consumer Finances (SCF), 1989-current
  - Good national and top representation, via oversample
  - e.g. Bricker, Henriques, Krimmel, Sabelhaus (2016)
- Estate tax data
  - Infer wealth from tax filing by modeling mortality
  - e.g. Kopczuk and Saez (2004), Johnson and Raub (2015)
- Income tax data (*inferred*)
  - Excellent coverage at top of income distribution, over time
  - No comparable *wealth* data
  - *Infer* wealth from income tax filings (Saez and Zucman 2016, Greenwood 1983)

Intro

- **Promising: measuring wealth at the top, over time**
- Surveys
  - SCF 1989-current
  - Good national and top representation (via oversample)
  - e.g. Bricker, Henriques, Krimmel, Sabelhaus (2016)
- Estate tax data
  - Infer wealth from tax filing by modeling mortality
  - Get the top but only small part of the top
  - e.g. Kopczuk and Saez (2004), Johnson and Raub (2015)
- **Income tax data (inferred)**
  - Excellent coverage **at top** of income distribution, **over time**
  - No administrative *wealth* data, though
  - *Impute* wealth from income tax filings (Saez and Zucman 2016, Greenwood 1983)

Intro

- A deeper look at capability to measure wealth with income tax data
  - Can these data identify those in the top?
  - How certain can we be of top share?
  - Can we characterize persistence over time?
  - Outside of the top?

Presentation Outline

1. Intro
2. Infer wealth
3. Noisy identification of the top?
4. Estimate persistence at the top?
5. Outside the top

Infer wealth using income tax data

- Use 4-year panel of administrative data derived from income tax returns (INSOLE)
  - $\approx$ 300,000 tax units *sampled* from master files
  - 2008-2011, centered on 2010

Infer wealth using income tax data

- Use 4-year panel of administrative data derived from income tax returns (INSOLE)
  - $\approx$ 300,000 tax units *sampled* from master files
  - 2008-2011, centered on 2010
- Capital income
  - Dividends, interest, capital gains, Sch E and C income...

Infer wealth using income tax data

- Use 4-year panel of administrative data derived from income tax returns (INSOLE)
  - $\approx$ 300,000 tax units *sampled* from master files
  - 2008-2011, centered on 2010
- Capital income
  - Dividends, interest, capital gains, Sch E and C income...
- Non-capital income
  - Wages, pension...

Infer wealth using income tax data

- Use 4-year panel of administrative data derived from income tax returns (INSOLE)
  - $\approx$ 300,000 tax units *sampled* from master files
  - 2008-2011, centered on 2010
- Capital income
  - Dividends, interest, capital gains, Sch E and C income...
- Non-capital income
  - Wages, pension...
- *Some* types of debt and taxes
  - Mortgage interest paid, property taxes paid...

Infer wealth using income tax data

- How to infer wealth? Capitalize income via **"gross capitalization"**

$$\hat{wealth}_i = \hat{house}_i + kg_i + \sum_{k=1}^{K} (income_i^k / return^k), K = 7$$

Infer wealth using income tax data

- How to infer wealth? Capitalize income via **"gross capitalization"**

$$\hat{wealth}_i = \hat{house}_i + kg_i + \sum_{k=1}^{K} (income_i^k / return^k), K = 7$$

- Used in SCF sampling
  - As one of two models (Kennickell 1998, Bricker + Henriques 2016)

Infer wealth using income tax data

- How to infer wealth? Capitalize income via **"gross capitalization"**

$$\hat{wealth}_i = \hat{house}_i + kg_i + \sum_{k=1}^{K} (income_i^k / return^k), K = 7$$

- Used in SCF sampling
  - As one of two models (Kennickell 1998, Bricker + Henriques 2016)
- Used to infer wealth
  - As in Greenwood (1983), Saez and Zucman (2016)

Infer wealth using income tax data

- How to infer wealth? Capitalize income via **"gross capitalization"**

$$\hat{wealth}_i = \hat{house}_i + kg_i + \sum_{k=1}^{K} (income_i^k / return^k), K = 7$$

- What are the correct model parameters and inputs?

Infer wealth using income tax data

- How to infer wealth? Capitalize income via **"gross capitalization"**

$$\hat{wealth}_i = \hat{house}_i + kg_i + \sum_{k=1}^{K} (income_i^k / return^k), K = 7$$

- What are the correct model parameters and inputs?
- Model parameters: **rate of return**?
  - Ratio: taxed income flow to FA asset stock? (Saez + Zucman 2016)
  - Market rates? (Greenwood 1983, Bricker + Henriques 2016)
  - Heterogeneous rates? (Kennickell 1998, Bricker + Henriques 2016)

Infer wealth using income tax data

- How to infer wealth? Capitalize income via **"gross capitalization"**

$$\hat{wealth}_i = \hat{house}_i + kg_i + \sum_{k=1}^{K} (income_i^k / return^k), K = 7$$

- What are the correct model parameters and inputs?
- Model parameters: **rate of return**?
  - Ratio: taxed income flow to FA asset stock? (Saez + Zucman 2016)
  - Market rates? (Greenwood 1983, Bricker + Henriques 2016)
  - Heterogeneous rates? (Kennickell 1998, Bricker + Henriques 2016)
- Model inputs: **income**
  - Annual income: permanent + transitory (Greenwood 1983, SZ 2016)
  - Permanent income: $perminc_i^k = (1/4) \sum_{t=1}^{4} income_{i,t}^k$

Infer wealth using income tax data

- Income: annual or permanent?
- Permanent
  - Wealthy realize income advantageous times (Wolfers 2015)
  - Do not want to allocate wealth that's not really there
  - Smooth income: $perminc_i^k = (1/4) \sum_{t=1}^{4} income_{i,t}^k$

Infer wealth using income tax data

- Income: annual or permanent?
- Permanent
  - Wealthy realize income advantageous times (Wolfers 2015)
  - Do not want to allocate wealth that's not really there
  - Smooth income: $perminc_i^k = (1/4) \sum_{t=1}^{4} income_{i,t}^k$
- Annual
  - Is household income a random walk (esp. capital income)?
  - No (DeBacker et al, 2012)

Presentation Outline

1. Intro
2. Infer wealth
3. Noisy identification of the top?
4. Estimate persistence at the top?
5. Outside the top

Identify wealthy using income data

Table: Predicted wealth percentiles in permanent vs. annual rankings

		2010 (annual income)				
		Bottom 95	90-95	Top 1	Top 0.1	Top 0.01
2008-11 (perm. income)	Bottom 95	0.99	0.16	0.00	0.00	0.00
	90-95	0.01	0.81	0.15	0.00	0.00
	Top 1	0.00	0.03	0.84	0.18	0.00
	Top 0.1	0.00	0.00	0.02	0.80	0.20
	Top 0.01	0.00	0.00	0.00	0.02	0.80

Note: categories in rows, columns are mutually exclusive groups

Identify wealthy using income data

Table: Predicted wealth percentiles in permanent vs. annual rankings

		2010 (annual income)				
		Bottom 95	90-95	Top 1	Top 0.1	Top 0.01
2008-11 (perm. income)	Bottom 95	0.99	0.16	0.00	0.00	0.00
	90-95	0.01	0.81	0.15	0.00	0.00
	Top 1	0.00	0.03	0.84	0.18	0.00
	Top 0.1	0.00	0.00	0.02	0.80	0.20
	Top 0.01	0.00	0.00	0.00	0.02	0.80

Note: categories in rows, columns are mutually exclusive groups

Identify wealthy using income data

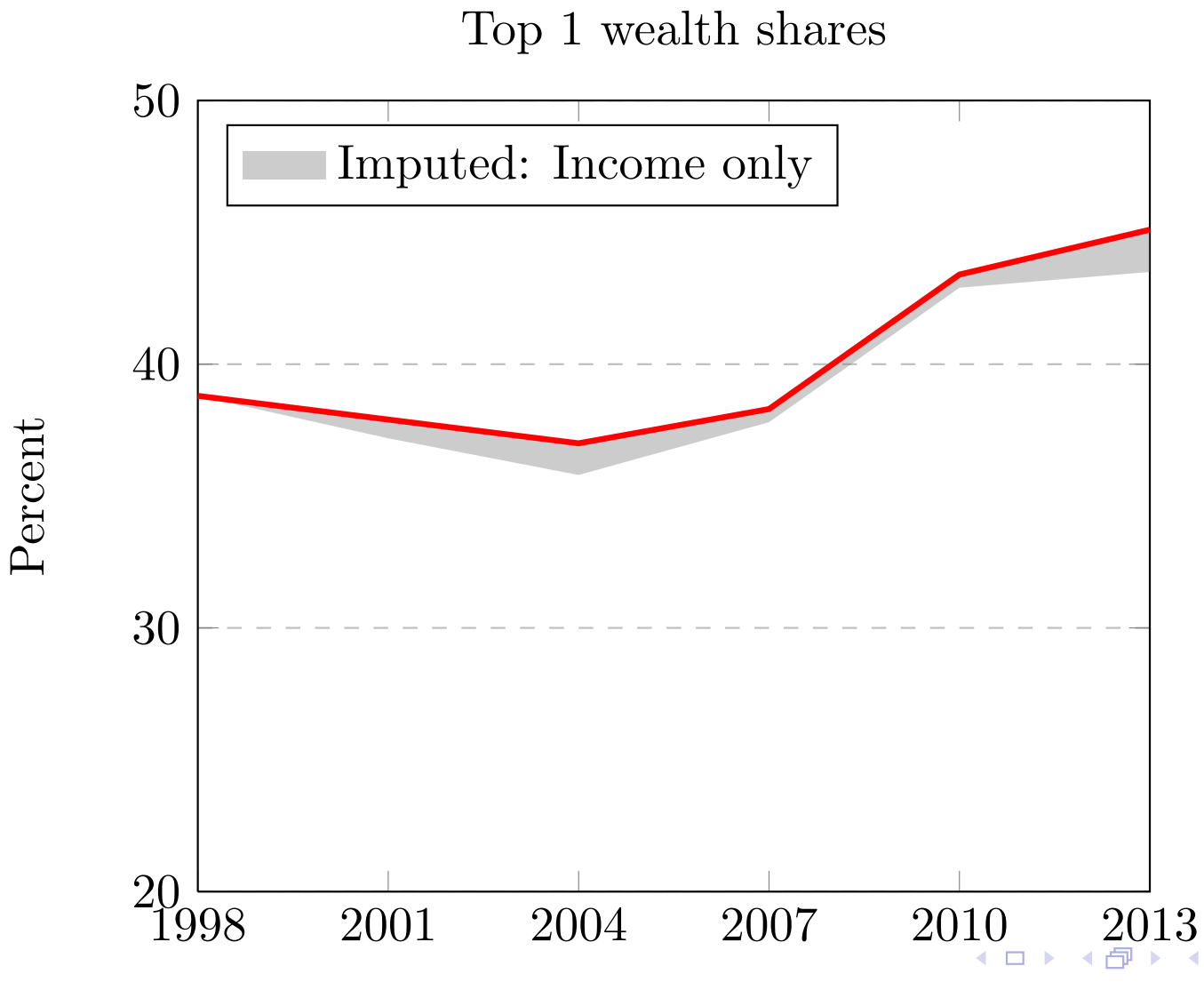
- Why? Transitory income changes
- If in top 0.1 pct in 2010 but not 2008-11: 50% larger income in 2010
- Mostly due to business and capital income

Imputed wealth uncertainty

- Estimate wealth via GC model (red line)
  - Use annual income and SZ rates of return
- Estimate uncertainty
  1. Income
     - Find Top 1 share using permanent income instead of annual income
  2. Rate of return
     - Find Top 1 share using market returns instead of SZ returns

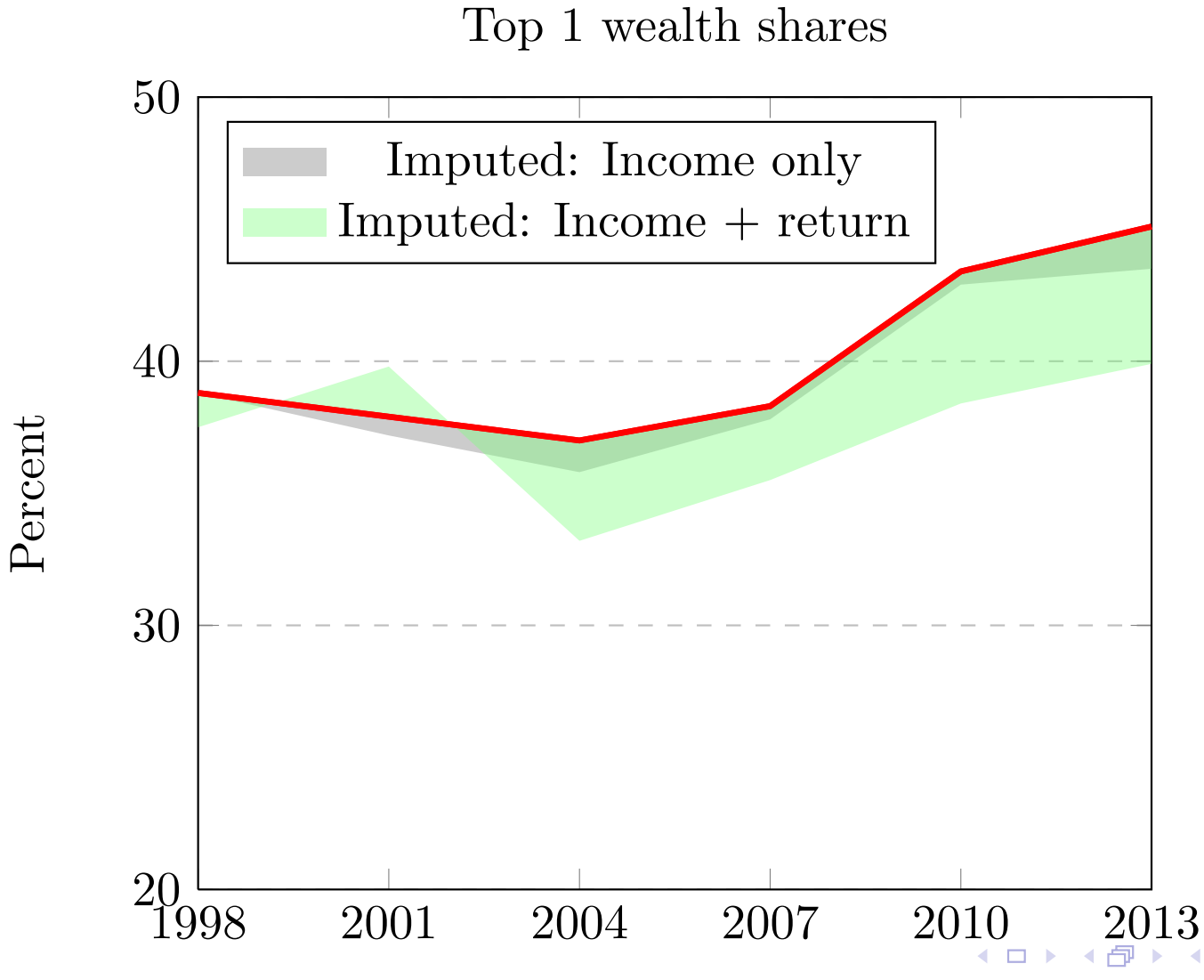
Imputed wealth uncertainty

- Income: annual vs permanent
-



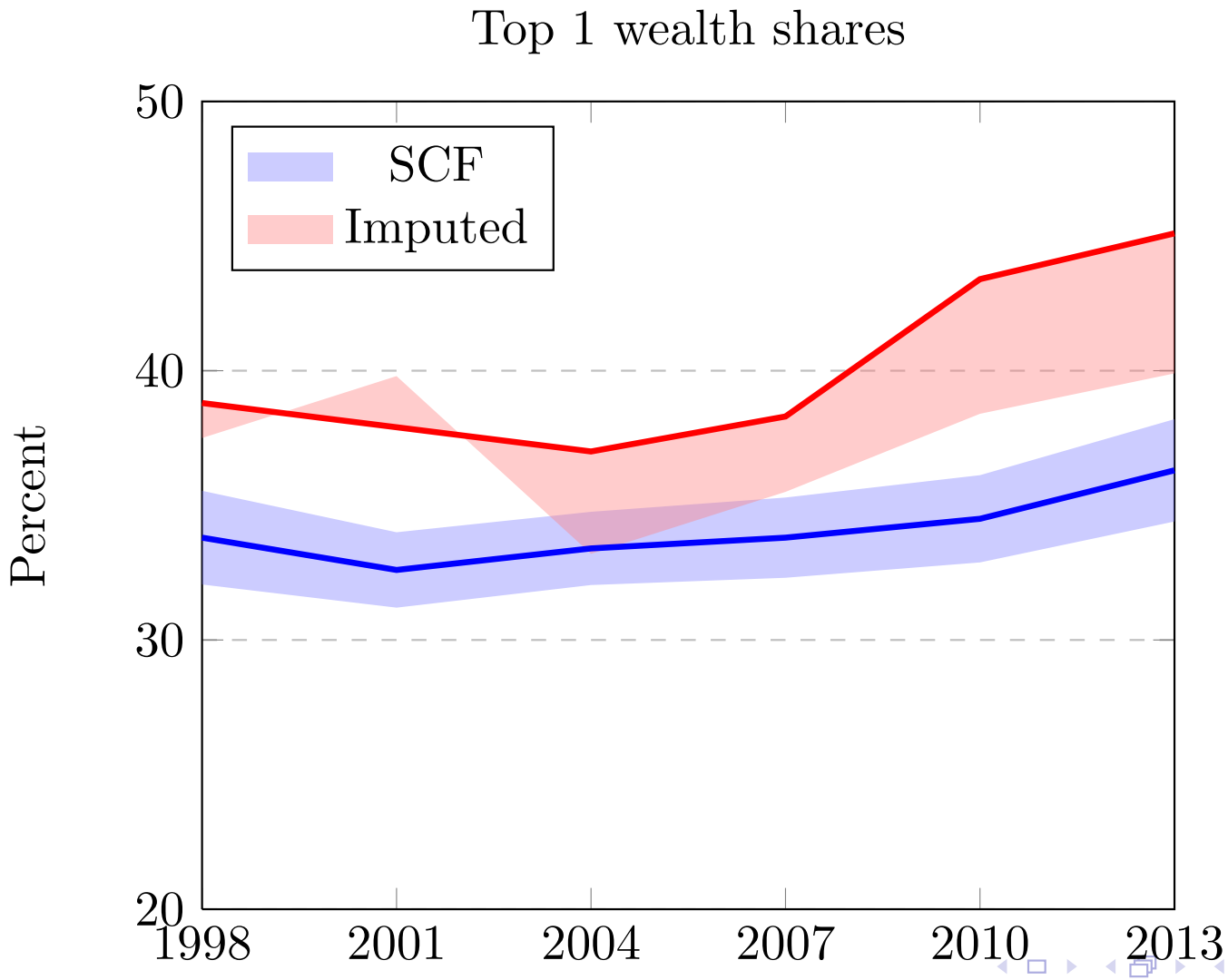
Imputed wealth uncertainty

- Income: annual vs permanent
- Returns: market vs. SZ



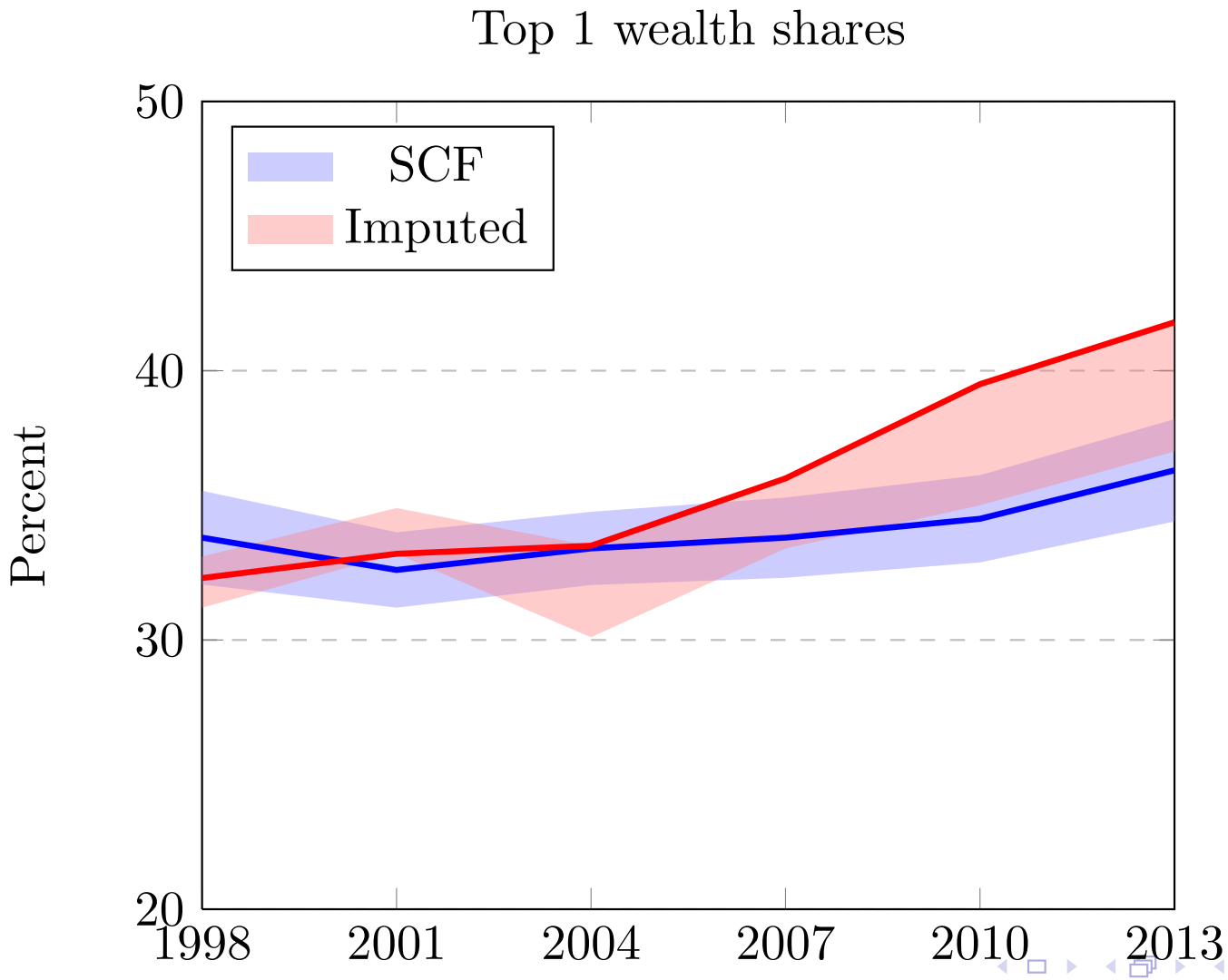
Imputed wealth uncertainty: GC vs. SCF

- Red area: modeling uncertainty
- Blue area: sampling + nonsampling uncertainty



Imputed wealth uncertainty: GC vs. SCF

- Red area: modeling uncertainty
- Blue area: sampling + nonsampling uncertainty



Presentation Outline

1. Intro
2. Infer wealth
3. Noisy identification of the top?
4. Estimate persistence at the top?
5. Outside the top

Persistence over time

- Membership in, e.g., Top 1 is noisily estimated with annual income
- Can get an *upper bound* on transitions with annual income

Persistence over time

Table: Persistence based on 1-year wealth estimates **(countinuously)**

Top 1

	2009	2010	2011
2008	0.80	0.71	0.64

Top 0.1

	2009	2010	2011
2008	0.73	0.62	0.55

Persistence over time: wealth vs. income

- Membership in top wealth groups more persistent than top income
- Income: $\approx$ 20 percentage points lower (Auten, Gee, Turner, 2013)

Presentation Outline

1. Intro
2. Infer wealth
3. Noisy identification of the top?
4. Estimate persistence at the top?
5. Outside the top

Outside of the top

- Membership in top groups noisily estimated
  - Small changes matter a lot
- Percentile changes small, in and out of top groups
  - average year-to-year change in top 1: 0.2 ppt
  - average year-to-year change out of top 1: 1.5 ppt

Conclusion

- Can these data identify those in the top?
  - Different answer for annual vs permanent income
- How certain can we be of top share?
  - Top shares *noisier* in income tax data despite excellent top coverage
  - Modelling: what income and what rate of return?
  - Heterogeneous returns: tighten?
- Can we characterize persistence over time?
  - Persistence greater than income
  - Upper bound on year-to-year transitions
- Outside of the top?
  - Overall rankings are fairly stable year-to-year