Inconsistencies in comparing relative prices over time: patterns and facts
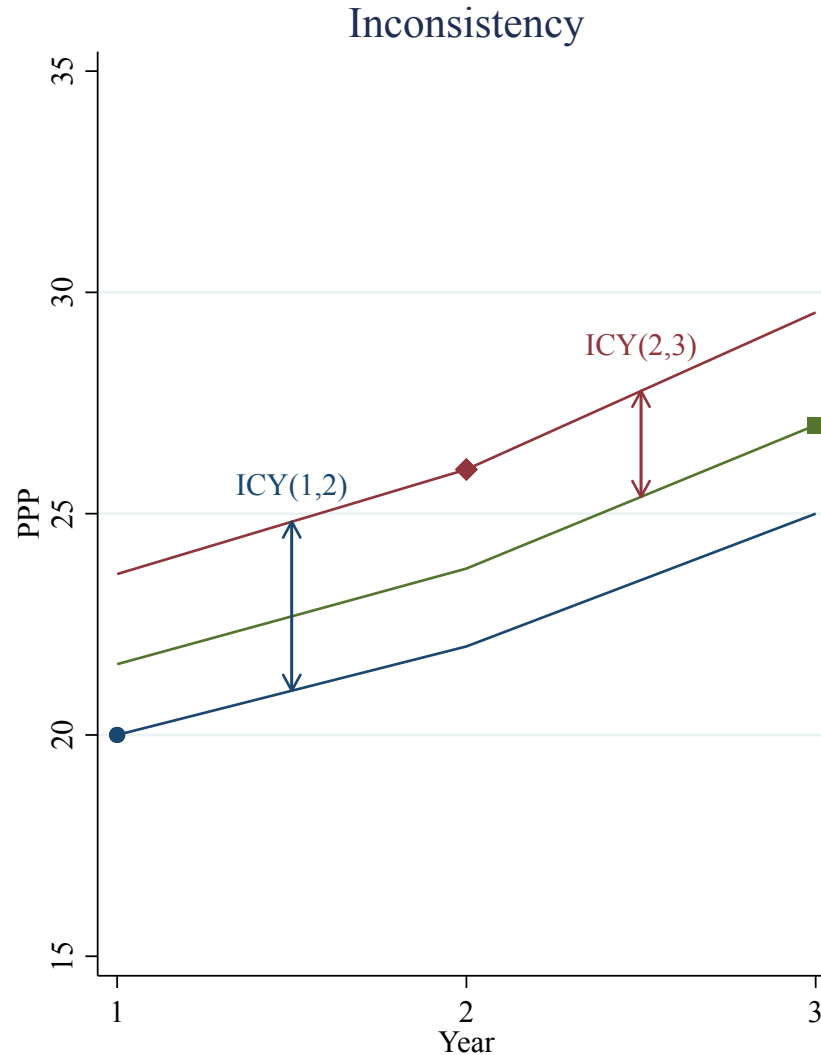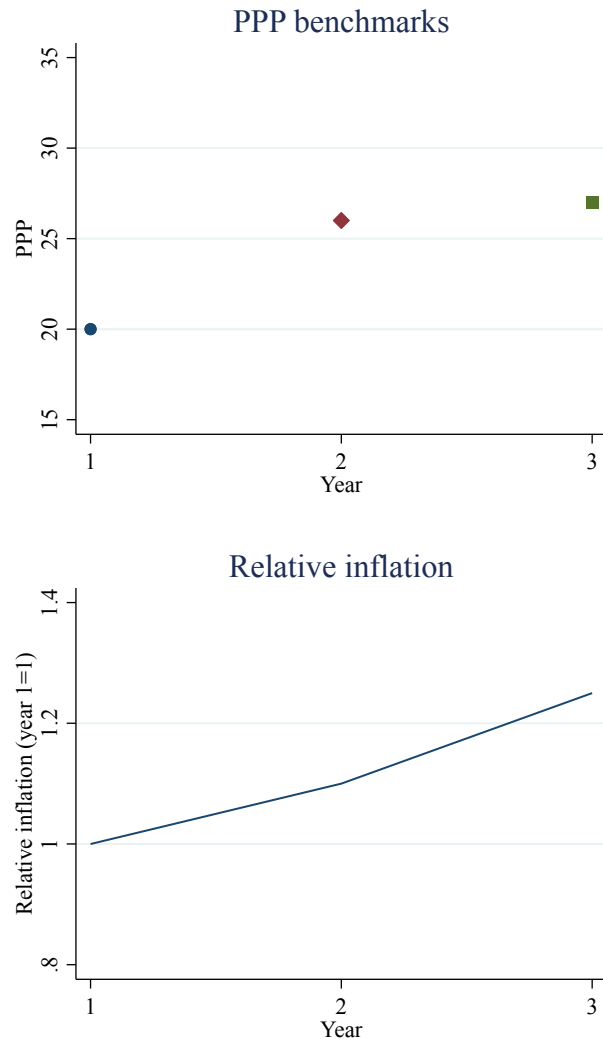
Robert Inklaar

Joint with Ryan Marapin, Jop Woltjer and Marcel Timmer

Meeting of the ICP TAG May 17–19, 2021

Inconsistency between benchmarks

2-country, 3-period example:
- Three PPP benchmarks
- Inflation in country A relative to country B
- Inconsistency as the difference between benchmarks and extrapolations implied by relative inflation

Background – ICP results

- Comparing ICP 2005 to ICP 2011 showed:
  - Large inconsistencies
  - More negative inconsistencies in countries with lower income levels => Extrapolating from ICP 2005 implied larger income inequality than ICP 2011

- Comparing ICP 2011 and ICP 2017 showed:
  - No systematic pattern of inconsistency by income level
  - Still large inconsistency in absolute terms

Background – Broader discussion

Deaton and Heston (2010):

• Comparing more disparate countries is generally harder

• Some products are harder to compare

Key questions

1. Are more recent global PPP benchmarks more consistent?

2. Is there more consistency between more similar countries?

3. Is there more consistency for products that are easier to price and compare across countries?

4. Would more frequent benchmark comparisons lead to more consistency?

5. Do inconsistencies distort the international income distribution?

Main answers

1. More recent ICP benchmarks are less inconsistent, suggesting improved measurement methods.
2. Price comparisons between countries:
    – With more similar expenditure patterns are less inconsistent.
    – With more similar income levels are (frequently) less inconsistent.
3. Inconsistency is not only high in hard-to-measure areas
4. When comparing PPPs across multiple benchmarks, there is no clear upward trend in inconsistency.
5. The only PPP benchmark where inconsistency varied systematically with income was ICP 1980.

Conceptual setting

- For individual items, the PPP change and relative inflation should match [e.g., in the sample of products with overlap between ICP and CPI]

- Inconsistency can be due to:
  - Weighting (Deaton & Aten, 2017):

$$\Delta \log PPP_{jk}^{vt} = \left(\Delta \log \pi_j^{vt} - \Delta \log \pi_k^{vt}\right) - \sum_i \frac{1}{2}\left(s_{ik} - s_{ij}\right)\left(\Delta \log \pi_{ij}^{vt} + \log \pi_{ik}^{vt}\right)$$

  - Differences in product sample and price measurement methods

Measurement approach

Inconsistency between the change in $PPPs$ and relative inflation, π from period v to t between country j and k:

$$d_{jk}^{vt} \equiv \Delta \log PPP_{jk}^{vt} - \left(\Delta \log \pi_j^{vt} - \Delta \log \pi_k^{vt} \right)$$

Summary measure (reference country independent):

$$RMSI^{vt} = \sqrt{\sum_j \left(d_{jk}^{vt} - \bar{d}_k^{vt} \right)^2}$$

With $\bar{d}_k^{vt} = \frac{1}{c} \sum_j d_{jk}^{vt}$

Empirical setting

- ICP 2011 and ICP 2017 plus:
  - Counterfactual ICP 2005 (adjusting for bias and methodology, Inklaar and Rao, 2017)
  - A 'pseudo global comparison' for 1996 constructed for PWT 6
  - Early ICP rounds: ICP 1970, 1975, 1980 and 1985
  - Based on PWT 10.0, so including export and import PPPs

- Basic-heading level PPPs for 2011 and 2017

Great expansion in country coverage

ICP Benchmark	Participating countries	vs. t-1	vs. t-2	vs. t-3	vs. t-4	vs. t-5	vs. t-6	vs. t-7
1970	16							
1975	33	16						
1980	60	27	14					
1985	63	41	26	14				
1996	115	59	51	30	13			
2005	145	99	55	53	32	16		
2011	177	142	110	63	59	32	16	
2017	175	173	140	109	63	58	32	16

Notes: Column 'vs. t-1' lists the number of countries that participated in both that comparison and the previous one, so 16 countries participated in ICP 1970 and those same 16 also participated in ICP 1975. Column 'vs. t-2' compares to two comparisons earlier, so only 14 of the countries that were in ICP 1970 were also in ICP 1980.

Consider full sample, but also balanced sample of 52 countries since ICP 1985
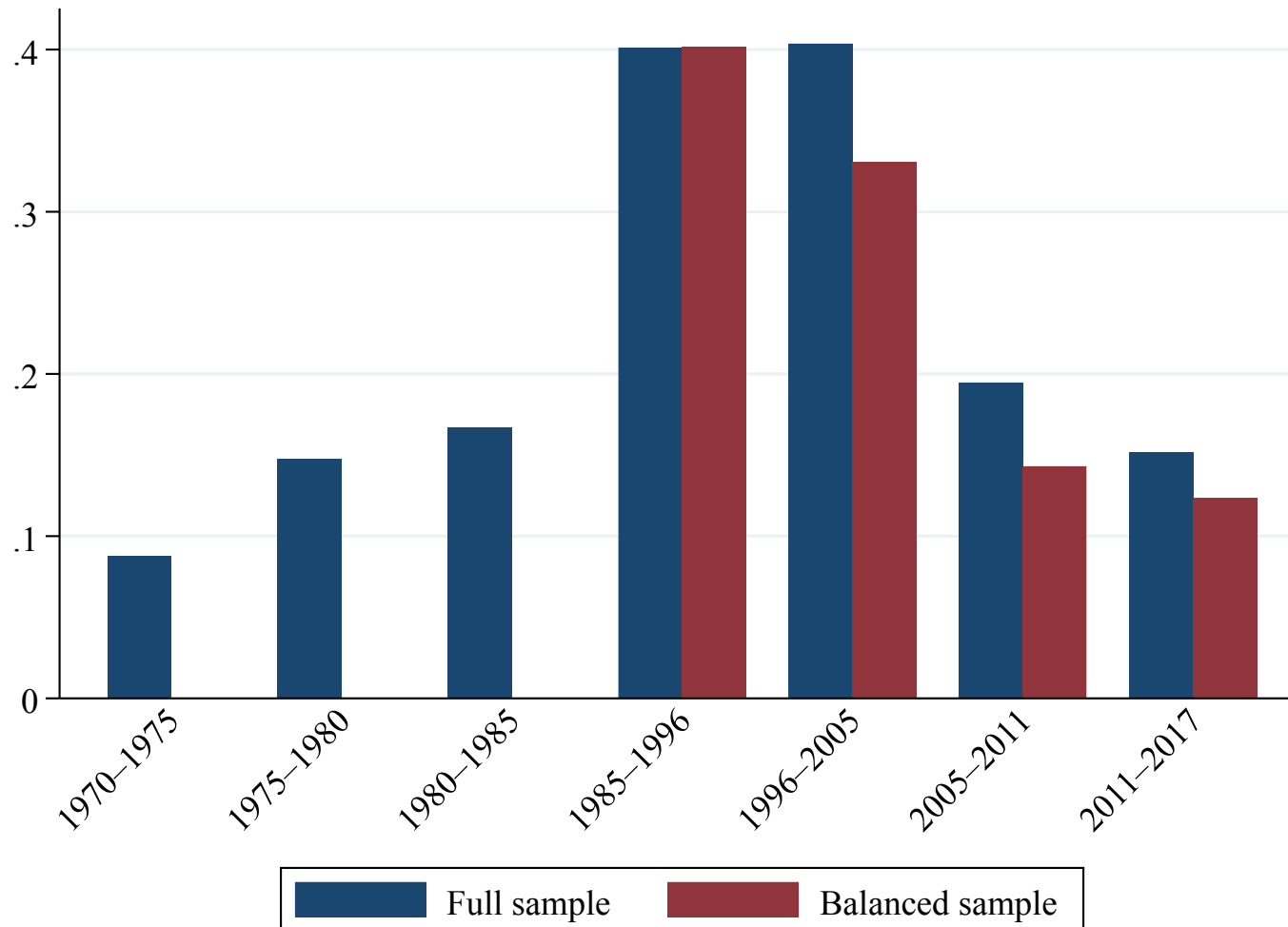
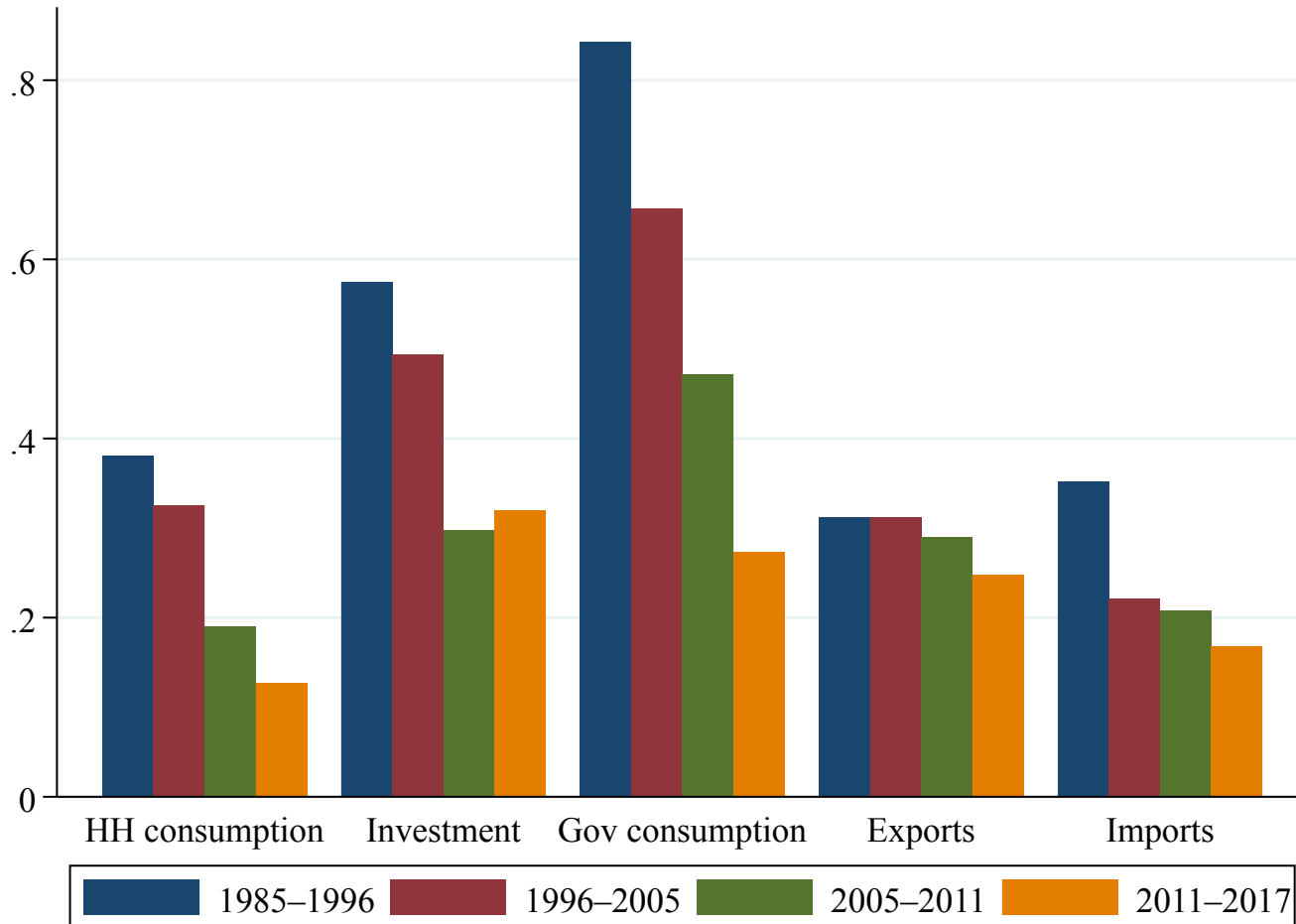# Reduced inconsistency for recent comparisons

Figure plots RMSI between consecutive PPP benchmarks

Reduced inconsistency mainly for ICP PPPs, not as clear for X&M PPPs

Comparing disparate countries is harder

Inconsistency d_{jk} is higher for countries with greater differences:

- In expenditure shares, $\delta s_{jk} = \sum_i (s_{ij} - s_{ik})^2$
- In income levels, $\delta y_{jk} = \sum_i (y_{ij} - y_{ik})^2$

|	GDP		Household consumption	
	Expenditure share	Income level	Expenditure share	Income level
1970–1975	0.45*	-0.04	0.44*	-0.06
1975–1980	0.52*	0.21*	0.41*	0.16*
1980–1985	0.27*	0.47*	0.11*	-0.03
1985–1996	0.22*	0.10*	0.20*	0.06*
1996–2005	0.22*	0.04*	0.07*	0.03*
2005–2011	0.21*	0.08*	0.23*	0.03*
2011–2017	0.20*	0.00	0.15*	0.10*

Reduction in inconsistency greatest in lower income quartiles
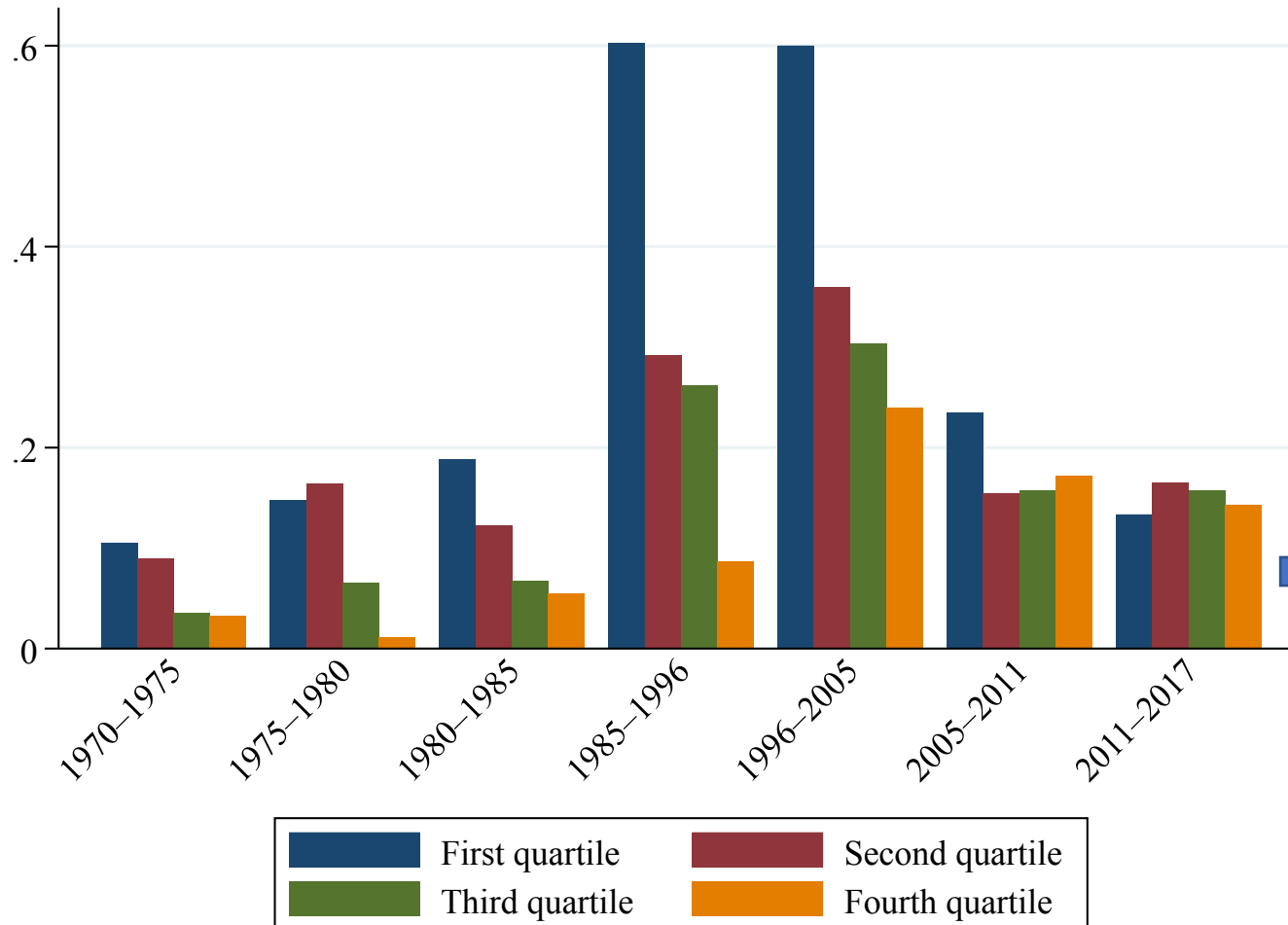
No notable difference across income quartiles comparing ICP 2011 and ICP 2017

Harder to measure consumption categories show higher RMSI
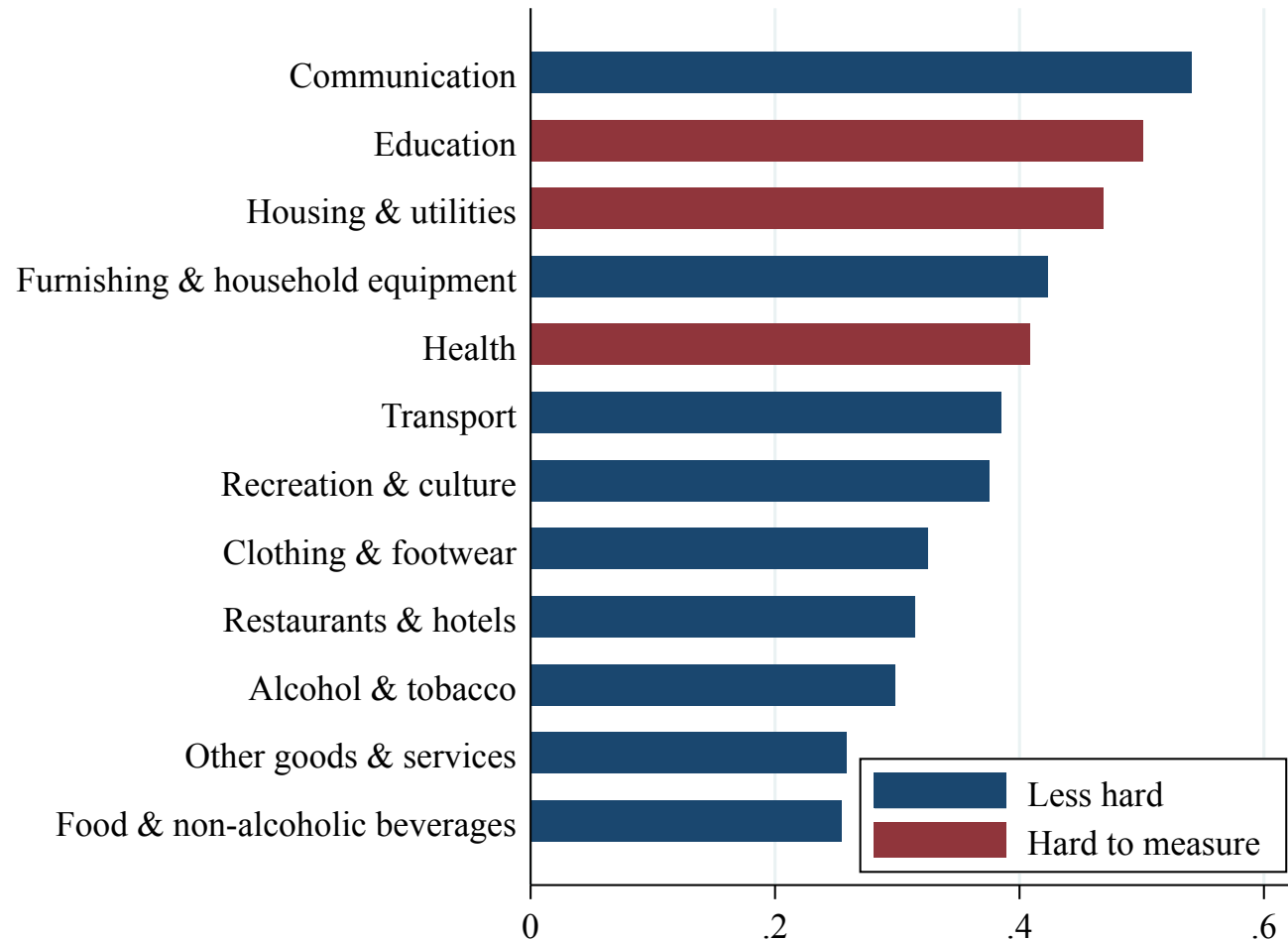
But:
a) 'Hard' vs. 'less hard' is very crude distinction
b) 'Less hard' categories also have high RMSIs

More time between benchmarks does not increase inconsistency

1. Start from a single benchmark and look down the diagonal

	Final benchmark:			
	1996	2005	2011	2017
1 benchmark apart (baseline)	0.40	0.33	0.14	0.12
2 benchmarks apart		0.29	0.37	0.17
3 benchmarks apart			0.32	0.35
4 benchmarks apart				0.33

No clear trend in inconsistency

university of
groningen

More time between benchmarks does not increase inconsistency

2. Look across the rows: no clear trend in inconsistency

	Final benchmark:			
	1996	2005	2011	2017
1 benchmark apart (baseline)	0.40	0.33	0.14	0.12
2 benchmarks apart		0.29	0.37	0.17
3 benchmarks apart			0.32	0.35
4 benchmarks apart				0.33

- In both cases, individual idiosyncrasies may drive the picture
- Is not a case against more frequent benchmarks, but may temper expectations

No recent shifts in international income distribution

- Do lower-income countries show systematically smaller or larger inconsistency?

  - Estimate: $d_{jk}^{vt} \equiv \alpha + \beta^{vt} \log y_j + \varepsilon_j$

Benchmarks	Coefficient	s.e.	# of countries
1970–1975	0.039	(0.019)	16
1975–1980	-0.092*	(0.021)	27
1980–1985	0.100*	(0.020)	41
1985–1996	-0.095	(0.054)	59
1996–2005	-0.033	(0.040)	99
2005–2011	-0.025	(0.014)	142
2011–2017	0.014	(0.010)	173

Conclusions

- Inconsistency is lower than before, …
- … but still substantial, with an RMSI on the order of 0.1–0.2
  - An adjustment in relative income levels of 10–20% is not uncommon
- Example: China/US

GDP level (bln)	ICP 2011	ICP 2017
China	$ 13,883	$ 19,617
United States	$ 15,543	$ 19,519
China/United States	89.3%	100.5%

- GDP level in ICP 2017 approximately equal, but given inconsistencies, how sure can we be that it was also not approximately equal in ICP 2011?