Distribution as Expenditure*

Alexis Antoniades
Georgetown University

ESCoE – London
March 13, 2018

* This research was made possible by support from an NPRP grant by the Qatar National Research Fund.

No Expenditure, No Weight, No Issue!

Big programs, big data:

- ICP, Billion Prices Project

Big ambitions:

- Measure cost of living across time and space, understand price behavior.

Big limitation:

- No information on quantities and/or expenditure.
- ICP: *"No weights reflecting the quantities of products sold are available so products are treated equally"* (same is true for studies using online data)

Questions

1. Do we introduce measurement error by using only prices?
2. Is such measurement error significant; can we quantify it?
3. Can we obtain expenditure information solely from prices?
4. Does the suggested approach reduce measurement error?

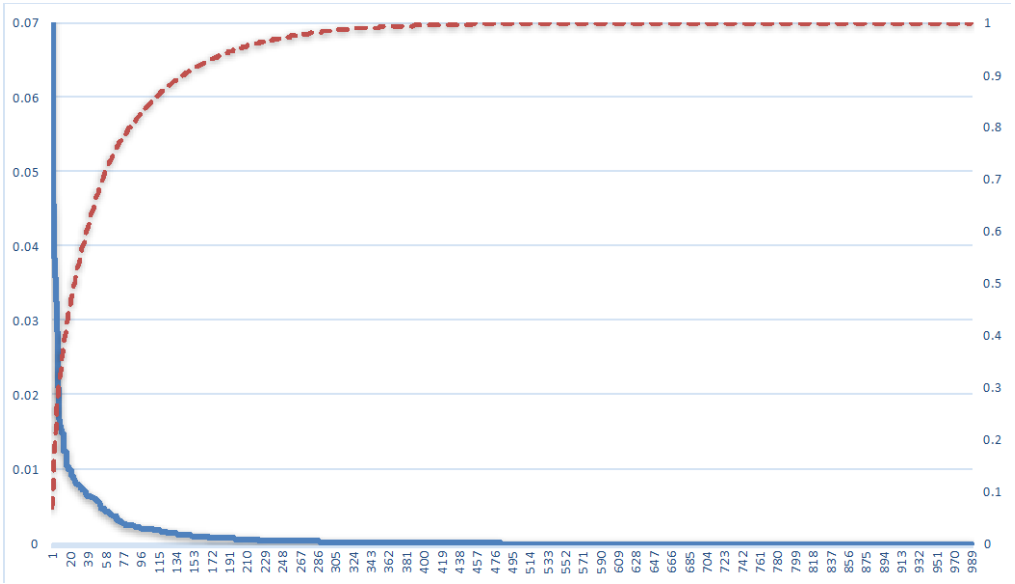
From $15 to $100 per bottle

VS

... to $20,000 - $1 million (or more)

A more *halal* example

- Dishwash, United Arab Emirates



Cumulative market share (right axis)

Market share (left axis)

- Not all products made equal: **Two distinct segments**
- Sales of top 2% = sales of lower 96% of products.
- Expect prices and price behavior to differ across segments.
- Need to be able to distinguish between products in the two segments.

Big Issue

- A few products account for the majority of sales.
- Assigning equal weights introduces measurement error.
- Measurement error is substantial (shown next)
- Need to weight products by importance. But how?
- For any innovation to work it should be easy to implement (in terms of cost and time) and should work with existing data (ICP, scraped online prices).

Outline

I. Propose an easy solution

- Obtain expenditure information solely from price observations
- Illustrate it

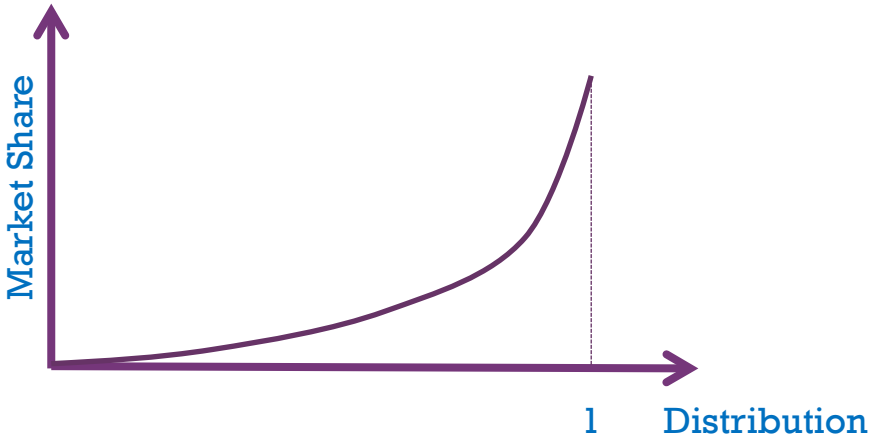
II. Quantify measurement error

1. Measure frequency and magnitude of price changes
   - Excl. weights understates freq by 25% (71%)
2. Measure inflation
   - Excl. weights understates 5-year infl. by 25% (73%)
3. Measure international price differences
   - Excl. weights overstates Δp by 200% (75%)
4. Measure exchange rate pass-through
   - Excl. weights understates long-run pass-through by half.

1 The Solution: Retail Distribution

Use distribution

- Distribution: number of outlets carrying a product over total outlets
- The relation between market share (MS) and retail distribution is convex: as distr ↑ MS ↑



1 The Solution: Retail Distribution

Use distribution

- Distribution is both a cause and a consequence of product market share.

1. **Distribution → Market Share:**
   - As distribution ↑ product availability ↑ => mkt share ↑
2. **Market Share → Distribution:**
   - In stores with limited shelf space owners have to decide which products to include => choose most popular (higher sales)

1 The Solution: Retail Distribution

- Convex relation is well documented in the marketing lit.

- Distribution measures in the literature:

$$\text{Numeric Distribution (\%)} = \frac{\text{Number of outlets carrying product}}{\text{Total number of outlets}}$$

$$\text{All Commodity Volume, ACV (\%)} = \frac{\text{Total sales of outlets carrying product}}{\text{Total sales of all outlets}}$$

$$\text{Product Category Volume, PCV (\%)} = \frac{\text{Total category sales of outlets carrying product}}{\text{Total category sales of all outlets}}$$

- 1st measure easiest to compute, does not take into account store size.
- Measures 2 and 3 take into account store size, have more data requirements (can't be computed in online or ICP data).

1 The Solution: Retail Distribution

Can we measure distribution? YES

- ICP: We already ask countries to audit *n* stores and give us the average price. All we have to do is ask in how many of the stores they found the product!
- Scraped online data: Can count the number of online sellers that provide a certain product.

1 The Solution: Retail Distribution

Can we measure distribution? YES

- We can even produce a weighted measure of distribution that takes into account the size of retailers based on total number of products sold (or in the case of the ICP, # of checkout counters, retailer type).

1 An Illustration: Convexity in Online P

Feenstra, Xu, and Antoniades (2016) - I

- Use online app to collect offline prices across Chinese cities in 2014.
- Multiple price per product across retailers in a city, no info on quantities
- Use data to compute numeric distribution (ND) and weighted (WD).

1 An Illustration: Convexity in Online P

City	Laundry Detergent		Personal Wash		Shampoo		Toothpaste	
	EANs	Retailers	EANs	Retailers	EANs	Retailers	EANs	Retailers
Beijing	929	11	1,273	11	1,041	10	1,024	11
Changsha	874	10	1,471	11	1,063	9	960	10
Chengdu	778	8	1,214	7	957	8	560	7
Chongqing	870	10	1,419	10	998	11	880	9
Dalian	661	6	986	4	775	5	655	3
Guangzhou	902	14	1,524	16	1,071	12	826	13
Hangzhou	805	8	1,210	8	975	8	788	8
Harbin	729	6	1,063	5	902	6	555	6
Hefei	968	10	1,325	10	1,090	9	1,069	8
Jinan	731	8	1,092	8	901	8	621	7
Kunming	579	5	978	5	773	5	422	5
Ningbo	676	7	1,074	8	842	7	569	7
Shanghai	999	12	1,456	12	1,226	10	1,032	12
Shenyang	929	10	1,383	10	1,084	11	847	10
Shenzhen	966	9	1,674	9	1,195	9	868	9
Suzhou	754	7	1,159	7	956	8	581	7
Tianjin	873	7	1,298	7	1,076	7	900	7
Wuhan	933	11	1,270	12	1,030	10	992	12
Wuxi	798	7	1,164	7	908	7	932	7
Xiamen	896	9	1,551	9	1,067	9	873	9
Xian	946	8	1,334	7	1,075	7	-	-

1 An Illustration: Convexity in Online P

Potential issues

- Number of audited stores small
- No info on outlet inclusion or exclusion
- Can we proxy retailer size with items available in the mob app? What does a missing item (price) mean? (not sold, or sold but price not recorded)

Implication:

- Produced measures of distribution will be noisy at best, meaningless at worse.

1 An Illustration: Convexity in Online P

Feenstra, Xu, and Antoniades (2016) – II

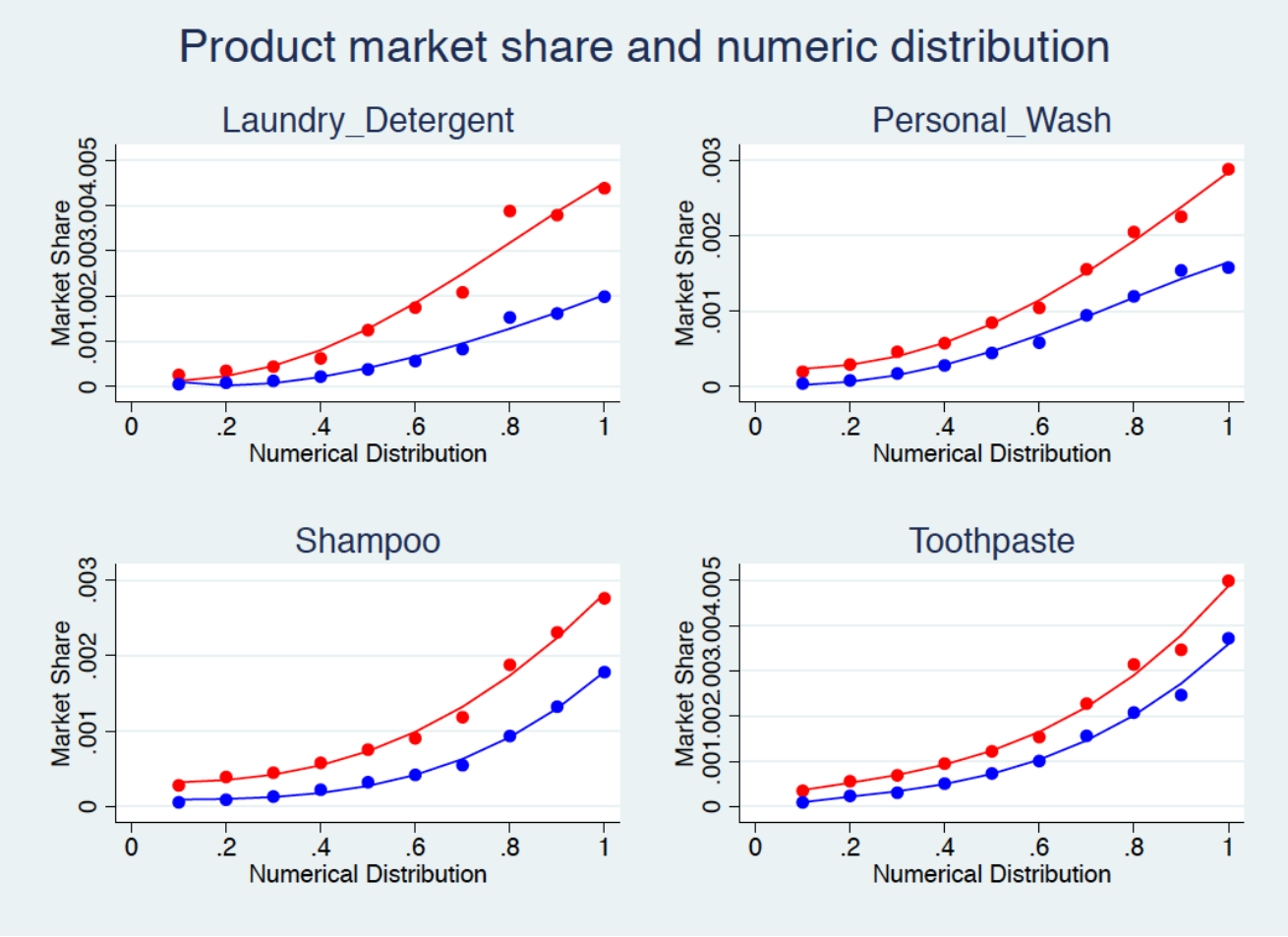
- Use Nielsen scanner data to collect prices and expenditure.
- Data averaged across retailers within a city => No info on distribution
- We can still obtain info on market shares

➢ Merge the two and plot retail distribution vs market share

- Allocate EANs into 10 bins based on distribution
- Report average or median market share for each bin
- Use either numeric distribution (ND) or weighted distribution (WD)

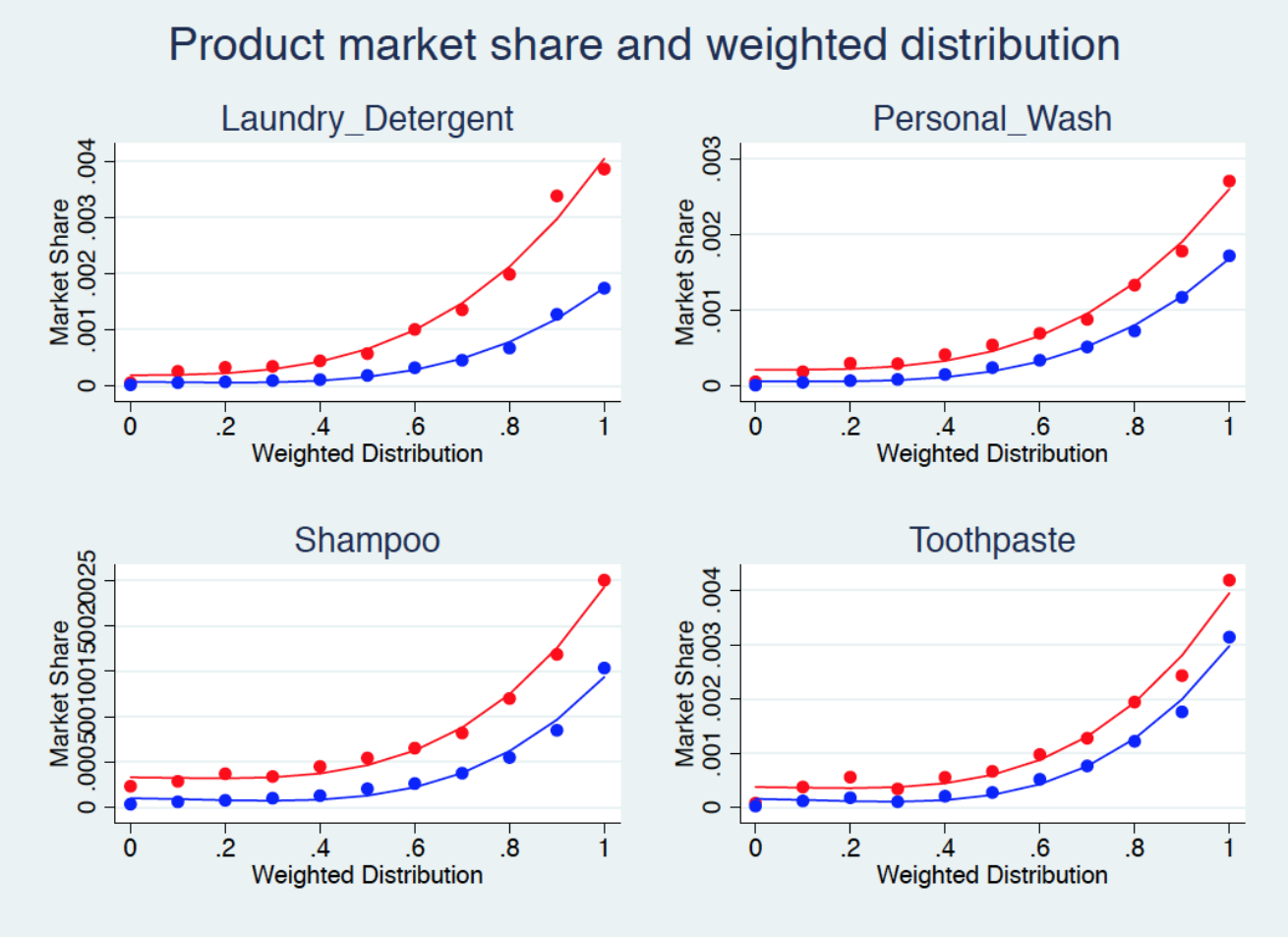
1 An Illustration: Convexity in Online P

Feenstra, Xu, and Antoniades (2016)



1 An Illustration: Convexity in Online P

Feenstra, Xu, and Antoniades (2016)



2 The Test

1. Can we measure retail distribution solely from prices? **YES**
2. Is measurement error large? How can we measure it?
3. Does approach work? How can we quantify this?

2 The Test

- Strategy: **Lab Experiment – Use scanner data to find out**

1. Use prices and quantities to measure or estimate:

   => Benchmark.

2. Use only prices, no quantities, compare with above:

   => Identify and quantify measurement error/bias.

3. Use prices & retail distribution as proxy for expenditure:

   => Quantify and evaluate performance of proposed methodology.

2 The Test

- Actually, four tests:

1. Measure price stickiness (freq and magnitude)
2. Measure inflation
3. Measure cross-country price differences
4. Measure exchange rate pass-through into retail prices

2 The Test

- Data
  - Nielsen barcode data
  - 30 product categories of FMCGs
  - Six countries (UAE, Oman, Bahrain, Kuwait, Qatar, Saudi Arabia)
  - 2006-2011
  - Data available *for each* outlet

2 The Test

(0) Prelim

- We predict market share from retail distribution using

(1) **mkt share = exp{-7.50 + 4.75*ND}** ND = numeric distribution

(2) **mkt share = exp{-7.90 + 114.7*WD}** WD = weighted distribution

where convexity coef. come from regressing ln(mktshare) on ND or WD

  - regression results

2 The Test

(0) Prelim

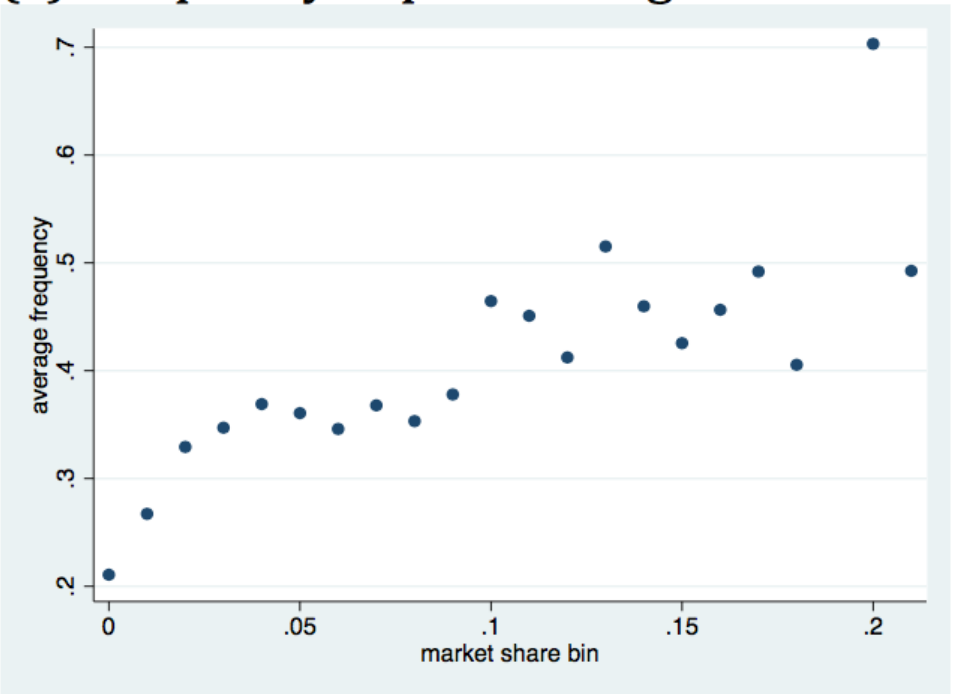
Table 2 - Market Share and Distribution Regression Results

	Dependent variable: ln(market share)			
	(1)	(2)	(1)	(2)
Numeric Distribution, ND	4.753***	4.818***		
	(0.000706)	(0.000702)		
Weighted Distribution, WD			114.7***	131.9***
			(0.0186)	(0.0185)
Country FE	NO	YES	NO	YES
Constant	-7.499***	-7.252***	-7.900***	-7.569***
	(0.000263)	(0.000796)	(0.000333)	(0.000795)
Observations	68,882,587	68,882,587	68,882,587	68,882,587
R-squared	0.397	0.410	0.356	0.429

2 The Test

(1) Frequency & magnitude of ΔP

(a) Frequency of price changes



- Most important items experience more frequent price changes.
- **Treating all products equally (i.e. excluding expenditure) will overstate the degree of price stickiness in the economy.**

2 The Test

(1) Frequency & magnitude of ΔP

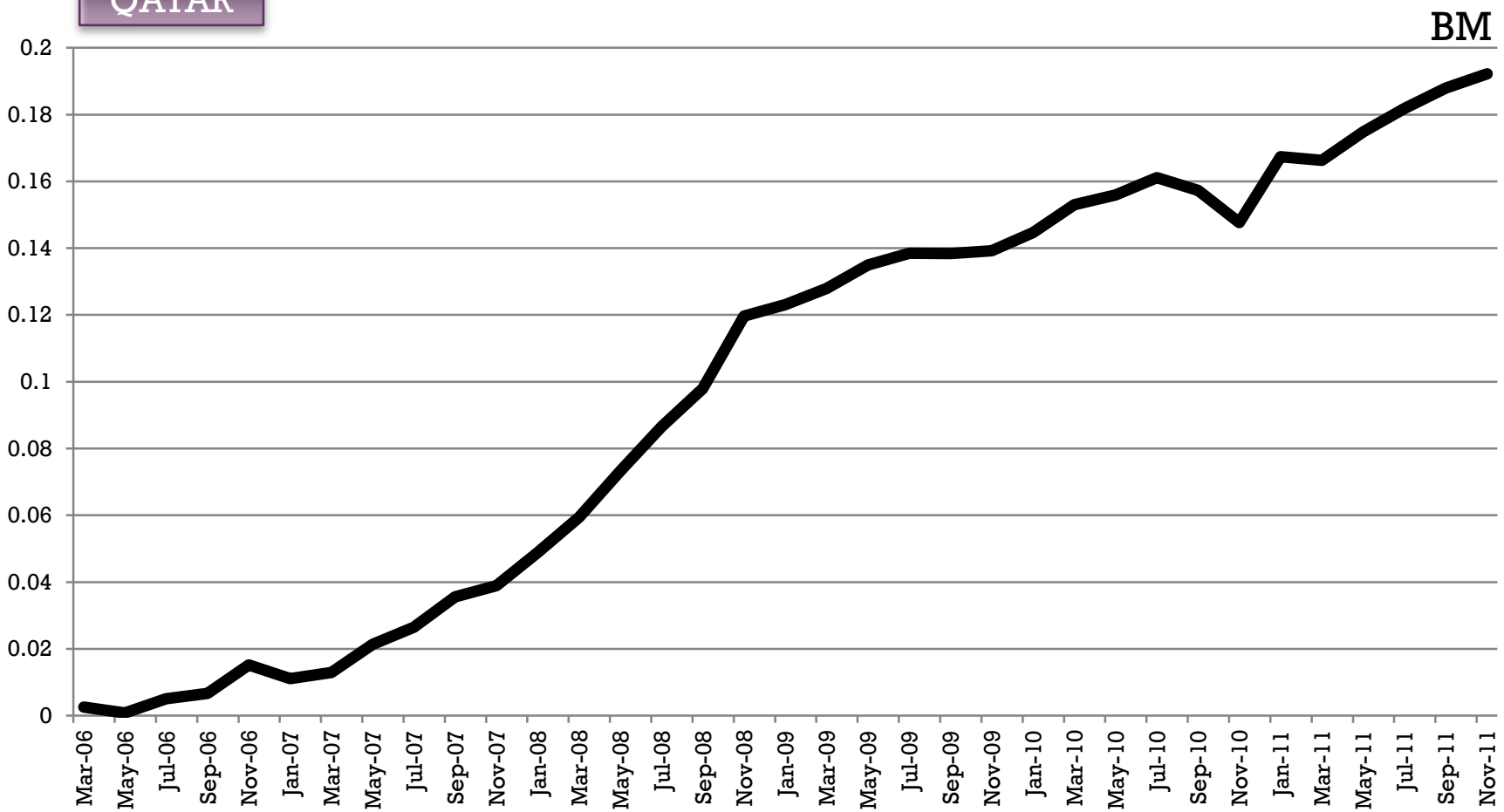
	Benchmark (P&Q)	P	P & NUM	P	P & NUM
	(1)	(2)	(3)	(4)	(5)
A. Frequency of Price Changes					
(i) With sales					
Continguous observations	0.30	0.23	0.27	-23%	-10%
Carrying regular price forward during sales and stockout	0.28	0.21	0.26	-25%	-7%
(ii) Without sales					
Continguous observations	0.24	0.19	0.22	-21%	-8%
Carrying regular price forward during sales and stockout	0.23	0.17	0.21	-26%	-9%

1. P, no Q : Measurement error is substantial (20%-26%), downward bias
2. P & ND/WD: Measurement error shrinks. It works.

2 The Test

(2) Inflation

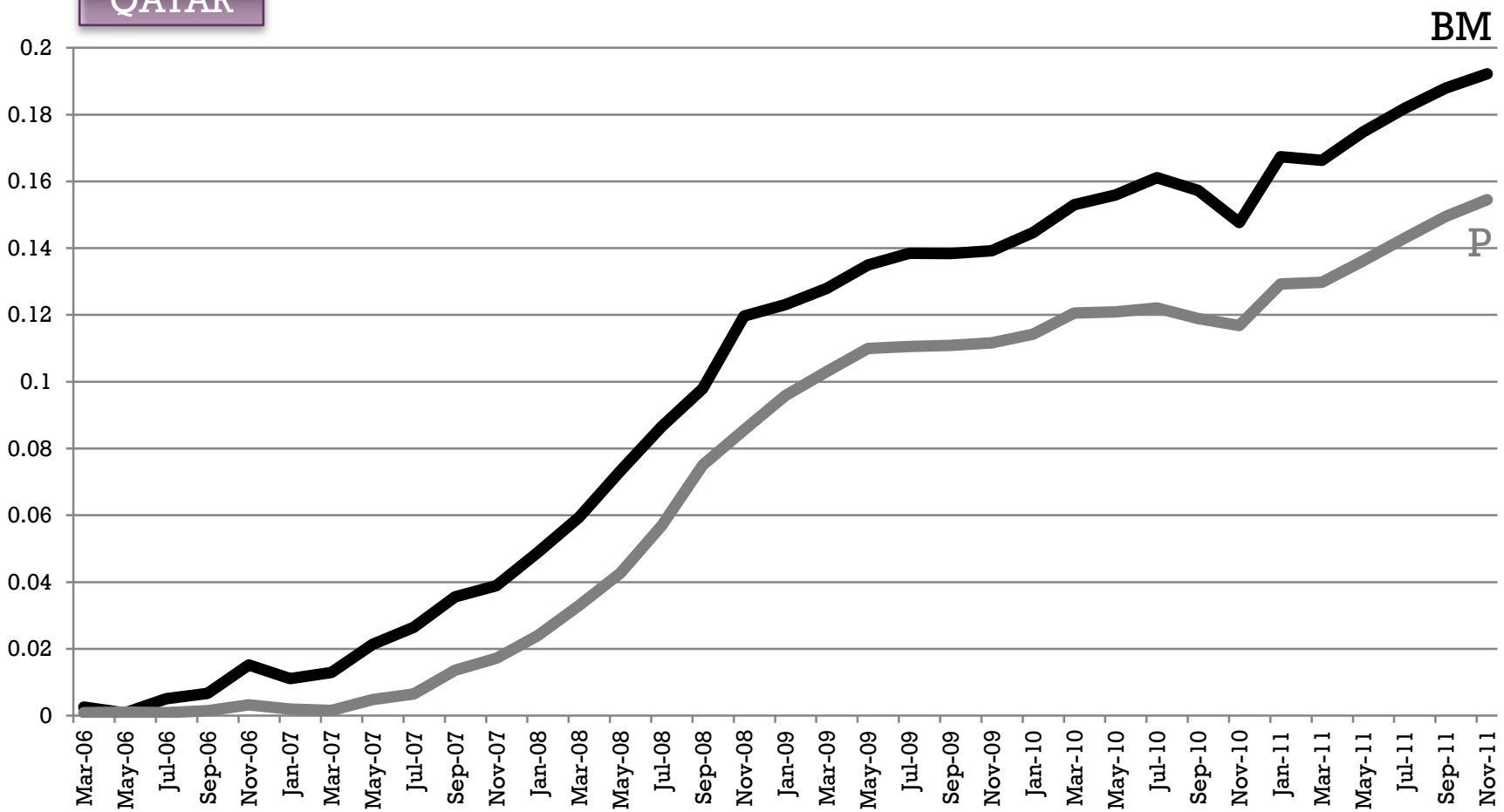
QATAR



2 The Test

(2) Inflation

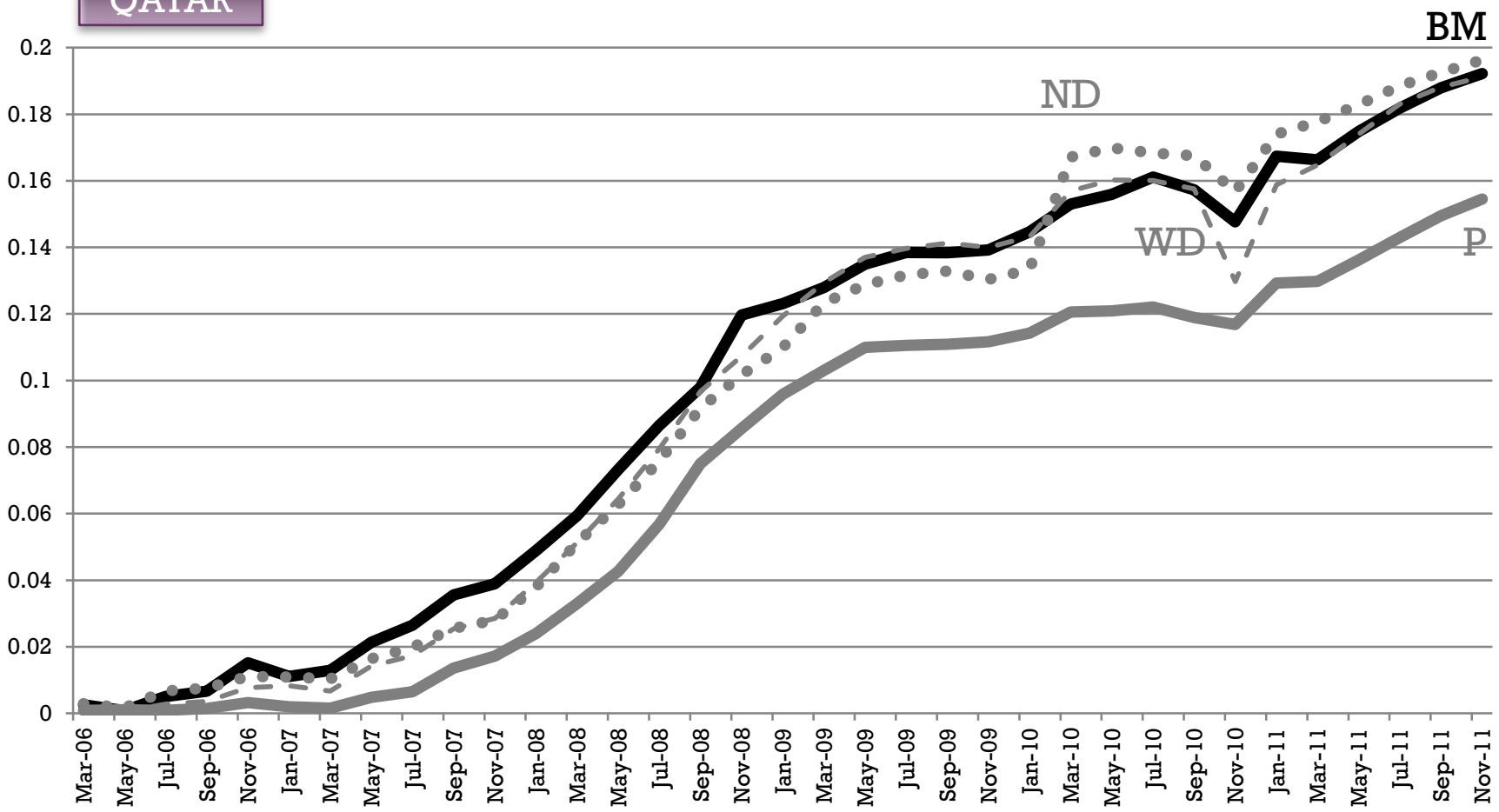
QATAR



2 The Test

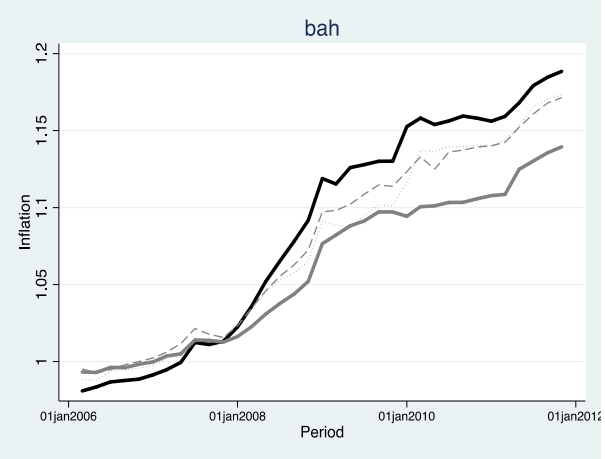
(2) Inflation

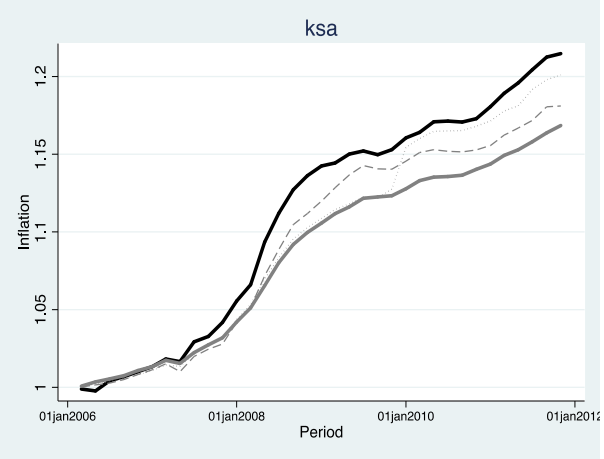
QATAR

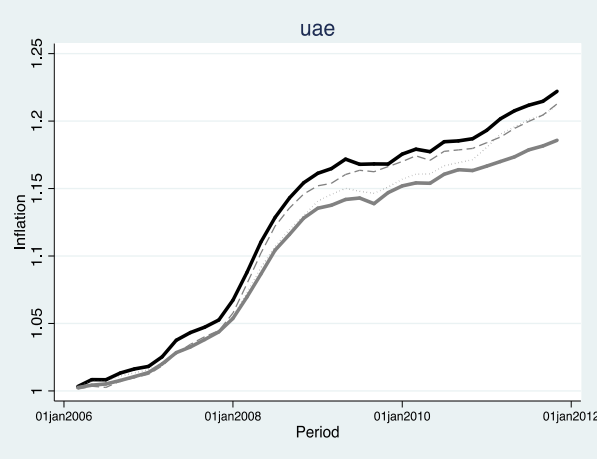


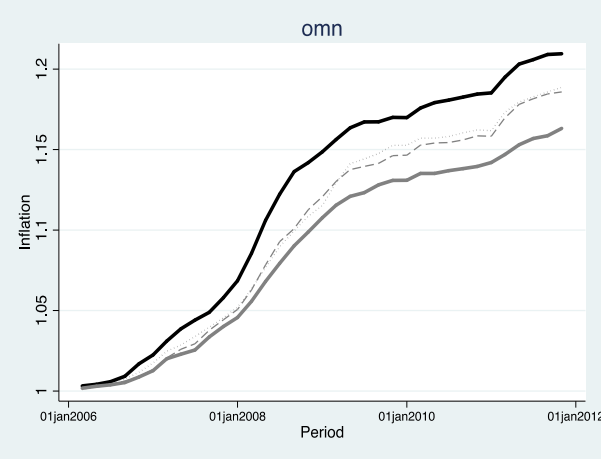
2 The Test

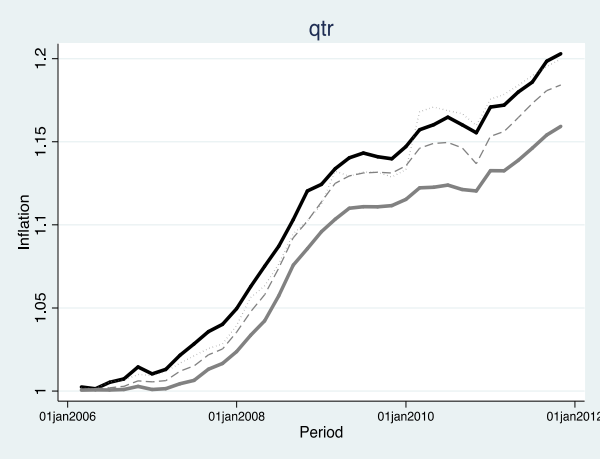
(2) Inflation

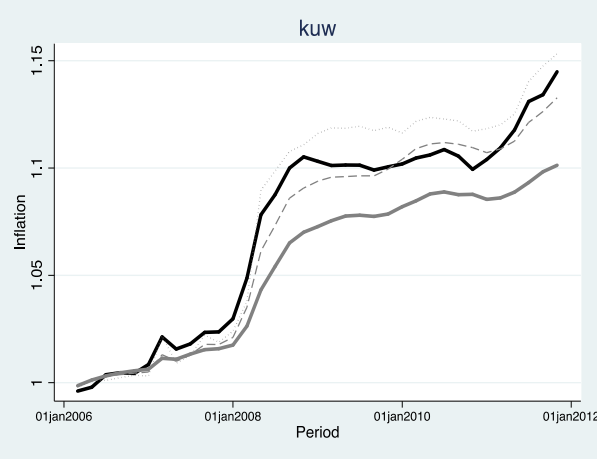












2 The Test

(2) Inflation

	Root Mean Square Error			Total Gap		
country	No WGT	ND	WD	No WGT	ND	WD
	(1)	(2)	(3)	(4)	(5)	(6)
Bahrain	0.13	0.03	0.01	-4.9	-1.5	-0.7
Saudi Arabia	0.08	0.03	0.02	-4.6	-1.4	-3.3
Kuwait	0.05	0.01	0.00	-4.4	0.8	-0.2
Oman	0.13	0.04	0.03	-4.6	-2.1	-1.3
Qatar	0.09	0.01	0.00	-4.4	-0.3	-0.6
UAE	0.05	0.02	0.00	-3.6	-1	0.1
Average	**0.08**	**0.02**	**0.01**	**-4.4**	**-0.9**	**-1**

Findings:

1. Democratic measure of inflation consistently understates the true cost-of-living.
2. Measurement error and bias shrink when ND or WD is used.

2 The Test

(3) International Price Differences & the ICP

Exercise: Simulated ICP (based on WB 2011 PPP confid. survey)

1. Manually allocate barcodes into PPP def's based on description
2. Audit *n* stores (n=10, 20, 50).
3. Randomly pick these stores; pref. for large stores.
4. Randomly pick an audit period within a year.
5. Pick the price. In case of multiple varieties, explore various rules (average, median, min, max, random, most import. item)
6. Compute average PPP difference across the GCC.
7. Repeat exercise 50 times (sort of bootstrapping the estimate).
8. Take average or median PPP diff across the 50 iterations.
9. Compare results for all versions (P&Q, P only, P and ND/WD)

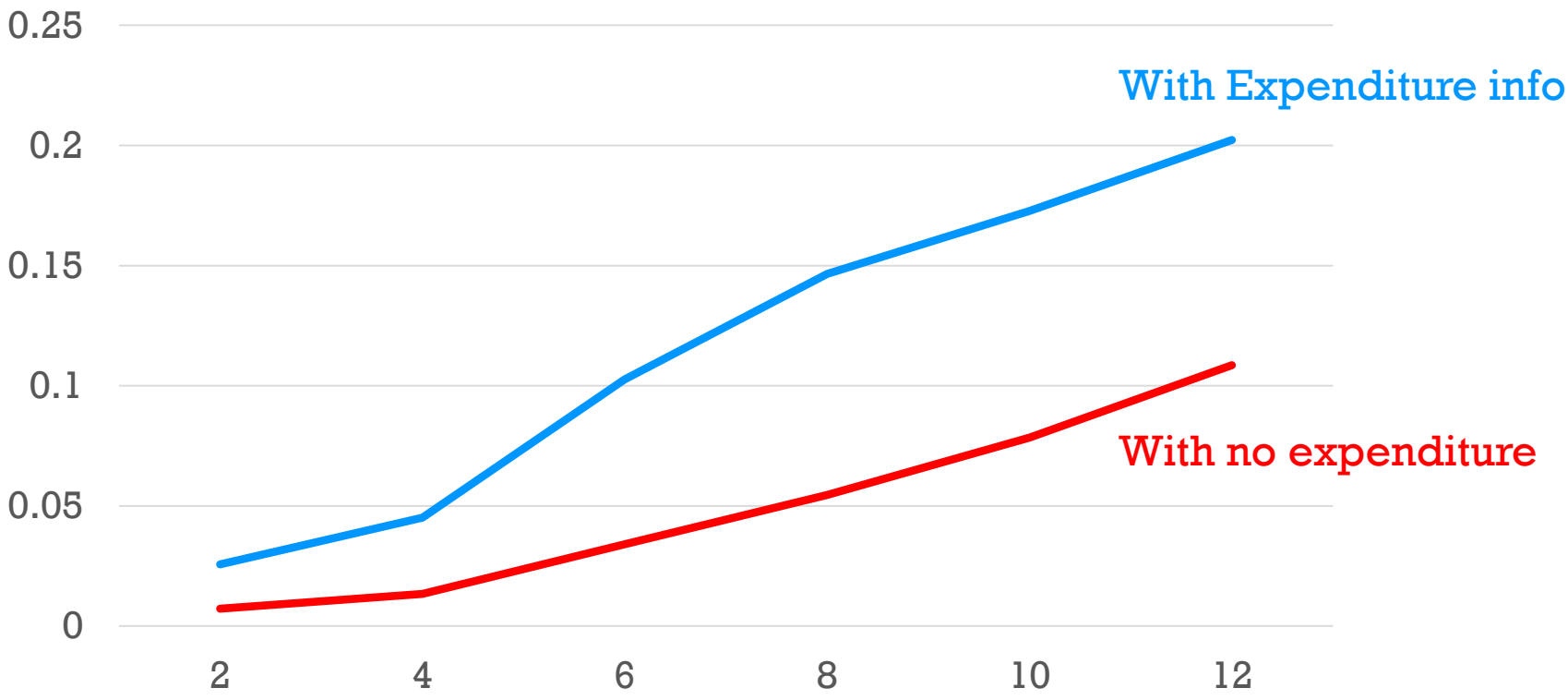
2 The Test

(3) International Price Differences & the ICP

Estimation Type	Outlets Audited	Average PPP difference among the GCC countries			
		Weights used			
		Expenditure (benchmark)	None	Distribution (ND)	Distribution (WD)
		(1)	(2)	(3)	(4)
avg	10	0.06	0.18	0.09	0.07
avg	20	0.05	0.17	0.06	0.06
avg	50	0.08	0.20	0.06	0.09
med	10	0.06	0.09	0.06	0.01
med	20	0.05	0.12	0.04	0.02
med	50	0.08	0.16	0.05	0.08
random	10	0.06	0.12	0.05	0.06
random	20	0.05	0.13	0.06	0.09
random	50	0.08	0.16	0.04	0.10

2 The Test

(4) Exchange Rate Pass-Through into Retail Prices*



* Antoniades and Zaniboni (2016) "Exchange Rate Pass-Through into Retail Prices," *International Economic Review*

3 Conclusion

- Availability of micro-price data offers great opportunities.
- But lack of quantities is a ***big issue ignored*** in recent studies.
- Our contribution:
  1. Document/Quantify how serious this issue is.
  2. Provide a simple, yet powerful methodological innovation to fix it.
  3. Evaluate/Quantify its performance.
- This simple innovation has great potential, lots of applications.