

Online and Official Price Indexes: Measuring Argentina's Inflation

Alberto Cavallo

April 2012

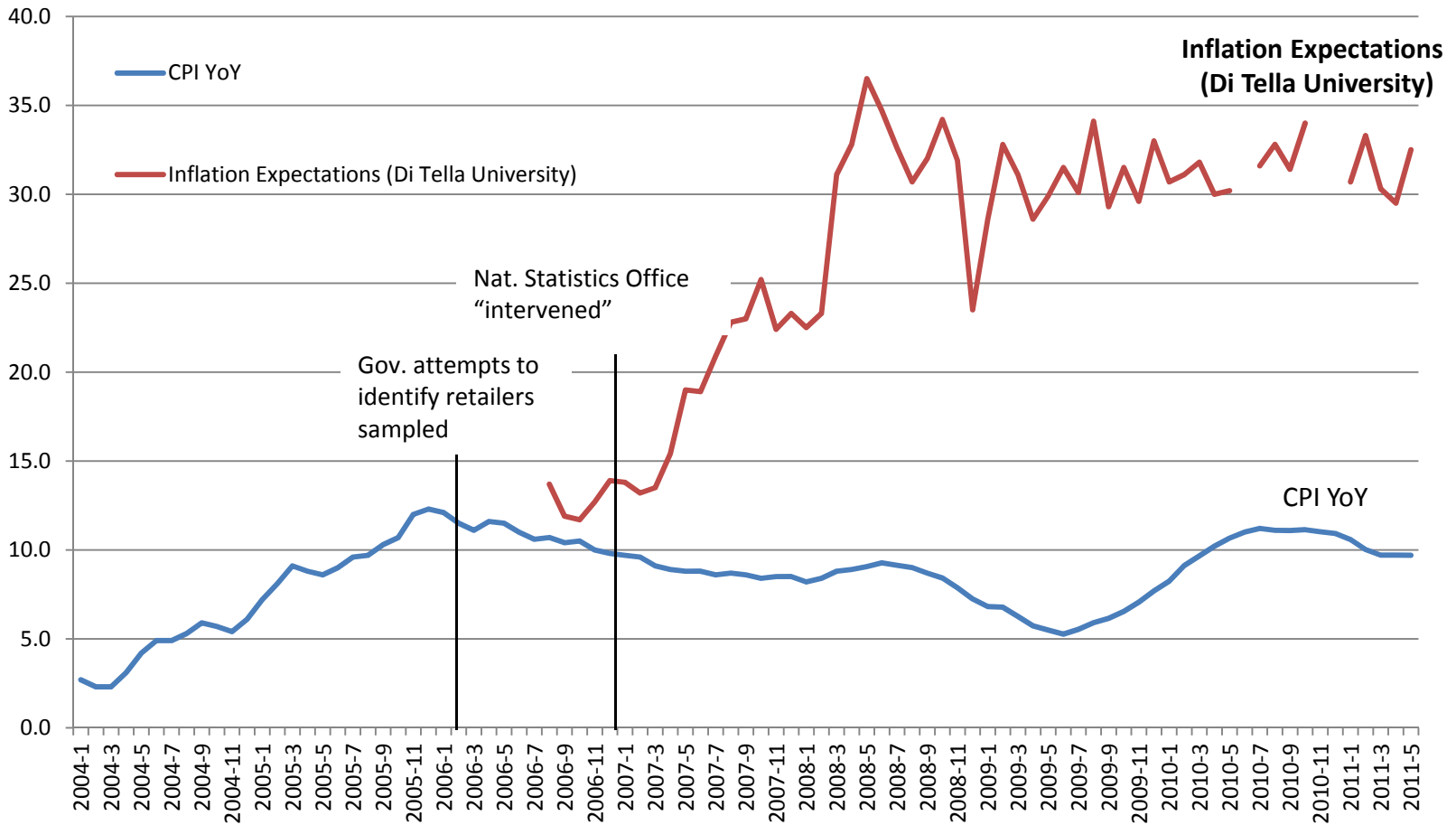
Motivation

- Online data can provide a daily product-level information to construct price indexes
- Advantages:
 - Cost
 - Speed & anticipation
 - Alternative series to complement official data:
 - Different methods, adjustments
 - Different aggregation levels and types
 - Alternative series in countries with “unreliable” data

Motivation

- Argentina's Inflation grew steadily until 2006, peaking at 12%.
- During 2006, employees of the National Statistics Institute (INDEC) reported attempts by the government to identify retailers being sampled
- In January 2007, the government intervened the INDEC and fired the people responsible for constructing the CPI
- Since then, INDEC has reported an annual inflation rate below 11%
- Widespread suspicion that these numbers are wrong:
 - INDEC announced “methodological changes”, but failed to disclose them
 - Some series stopped being reported (eg. National CPI)
 - Employees of INDEC denounced a manipulation of price series
 - Economists and provincial governments began to report higher inflation rates (currently between 25%-30%)
 - Pro-government Unions get salary increases of 25-30%
 - Official data → increasingly disconnected from consumer's inflation expectations

CPI and Inflation Expectations



Recent News

THE WALL STREET JOURNAL.

11/4/2010

BUSINESS | November 24, 2010

Argentina Asks IMF For Help on Data

4/5/2011

BUSINESS | April 5, 2011, 6:11 p.m. ET

IMF to Push Argentina for New Price Index

4/12/2011

BUSINESS | APRIL 12, 2011, 5:23 P.M. ET

Argentina Fines More Economists Over Inflation Estimates

7/9/2011

LATIN AMERICA NEWS | JULY 9, 2011

Argentina Charges Economists

2/1/2012

Bloomberg

IMF Criticizes Argentina's 'Absence of Progress' on Inflation, GDP Data

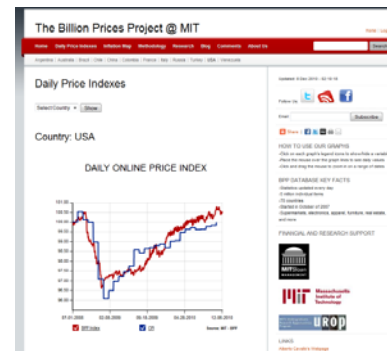
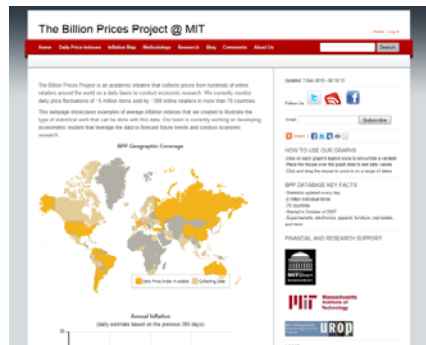
By Eliana Raszewski - Feb 1, 2012 10:01 PM ET

This Paper

- Builds alternative price indexes using daily online data from 5 Latin American countries (2007 to 2011) , using official CPI weights and methods.
- Shows that online price indexes can track official statistics in Brazil, Chile, Colombia, and Venezuela.
- Not in Argentina:
 - Inflation is between 2 or 3 times higher than officially reported
 - Accumulated difference over 3.5 years is 65 percentage points → huge implications for poverty, gdp
- Differences are not driven by data or methodology

Data: Billion Prices Project

- Joint with Roberto Rigobon
- “Scraped Data” → software that scans the code of web-pages and records all available price information
 - Daily sampling frequency
 - In 5 main categories of goods: supermarkets , electronics, furniture, apparel, real estate
 - Currently more than 1500 retailers in over 70 countries



What is Scraped Data?

- Scraped data is collected from public online sources
- A software downloads a webpage, analyses the html code, collects prices, and stores it in a database

BIENVENIDA PASO1 PASO2 PASO3 Block

Buscar

Buscar múltiples productos

Resultado Búsqueda

> LECHE CONDENSADA

- Uniformes Escolares
- Librería
- Almacén
- Aseo y Limpieza
- Perfumería y Farmacia
- Bajos en Calorías
- Bebés y Niños
- Bebidas y Jugos
- Carnicería
- Congelados
- Cóctel
- Ferretería
- Fiambrería
- Frutas y Verduras

Lácteos/Leche Condensada/Evaporada

Buscar por marca ... Ordenar por Precio

Producto	Descripción	Precio	Cantidad	Comprar
	Leche Condensada Nestlé Pack 3 unidades, Lata 200 grs. c/u \$xLitro:\$1.998	\$1.199 Uni	<input type="text" value="1"/> Uni	<input type="button" value="agregar"/>
	Leche Evaporada Ideal Lata 400 grs. \$xKilo:\$2.473	\$989 Uni	<input type="text" value="1"/> Uni	<input type="button" value="agregar"/>
	Leche Evaporada Jumbo Lata 410 grs. \$xKilo:\$2.193	\$899 Uni	<input type="text" value="1"/> Uni	<input type="button" value="agregar"/>
	Leche Condensada Nestlé Envase flexible 350 grs. \$xKilo:\$2.569	\$899 Uni	<input type="text" value="1"/> Uni	<input type="button" value="agregar"/>
	Leche Condensada Nestlé Descremada, Lata 395 grs. \$xKilo:\$2.023	\$799 Uni	<input type="text" value="1"/> Uni	<input type="button" value="agregar"/>

	_ID	_ID2	_PRODUCTO	_MARCA	_TAMANO	_BULKPRICE	_PRICE
1	3429	266235-ST	Leche Condensada	Leche Sur	Lata 395 grs.	xKilo:\$1.744	689
2	3422	266231-ST	Leche Condensada	Nestlé	Descremada, Lata 395 grs.	xKilo:\$2.023	799
3	995	619436-ST	Leche Condensada	Nestlé	Envase flexible 350 grs.	xKilo:\$2.569	899
4	3804	399781-ST	Leche Condensada	Nestlé	Lata 397 grs.	xKilo:\$1.761	699
5	11676	668674-ST	Leche Condensada	Nestlé	Pack 3 unidades, Lata 200 grs. c/u	xLitro:\$1.998	1.199

Compared to other Data Sources

Table 1: Alternative Data Sources

	Scraped Data	CPI Data	Scanner Data
Data Frequency	Daily	Monthly - Bi-Monthly	Weekly
Countries Available for Research	~50*	10-15	<5
All Products in Retailer (Census)	Yes	No	No
Comparable data across countries	Yes	Limited	Limited
Details: sale, price control, other	Yes	Limited	Yes
No <i>Forced</i> Substitutions	Yes	No	Yes
Real-Time data availability	Yes	No	No
Product Categories Covered	Few	Many	Few
Retailers Covered	Few	Many	Few
Quantities Sold	No	No	Yes

Notes: *Data from over 70 countries are currently being collected by the Billion Prices Project (bpp.mit.edu). ** Only goods purchased are scanned.

Methodology

- Use data from the largest supermarket in each country
- CPI Categories included: food , beverages, and household products

Table 1: Data Description

	Argentina Retailer #1	Argentina Retailer #2*	Brazil	Chile	Colombia	Venezuela
Starts	10/7/2007	23/7/2007	10/10/2007	10/24/2007	11/13/2007	04/16/2008
Ends	3/24/2011	03/20/2011	03/01/2010	03/20/2011	03/24/2011	03/01/2010
Categories	Yes	No	Yes	Yes	Yes	No
Prices P/day (mean)	11,560	4,790	11,000	12,000	5,000	9,256
Total Products	26,333	10,929	21,804	3,5432	9,166	20,847
Price Changes	204,449	136,781	25,9875	12,0112	76,979	94,808
Market Share (Supermarkets) **	28%	n/a	15%	27%	30%	n/a

Note: *Argentina's Retailer #1 is used in all online prices indexes reported in this paper for that country. Argentina's Retailer #2 is used only in the robustness results shown in Figure 5. **Market shares are based on claims posted on the corporate webpages of each supermarket's parent companies.

Methodology

- Price indexes:
 - Daily
 - Include sales
 - No product substitutions
 - Use all products available in each retailer
 - Missing values *within* price spells are completed using the last available price for each product.

Methodology

- Combine with standard CPI methods in these countries:
 1. price changes are obtained at the product level,
 2. then averaged inside categories using geometric means,
 3. then aggregated across categories with a weighted arithmetic mean
- No special adjustments or corrections (eg: quality)
- Use official category weights

Methodology

- I use official CPI category weights and compare online series to an equivalent official index (covering food, beverages, and household products)



home > seções > b

Cerveja

Item	Orde
	Cervi
	Cerve
	Cerve

URL: www.retailer.com/?cat=cerveja

Table 2: CPI Category Weights (%)

Category Name	Argentina	Brazil	Chile	Colombia
Baby Food			0.08	
Beer At Home	0.27	0.30	0.36	0.43
Distilled Spirits At Home	0.11		0.32	0.20
Wine At Home	0.48		0.37	0.50
Cigarettes		0.83	0.95	0.12
Dental & Nonelectric Shaving Products	0.29	0.16	0.25	0.40
Deodorant/Suntan Preparations	0.29	0.45	0.17	
Electric Personal Care Appliances			0.05	
Cosmetics Nail Preparations & Implements	0.63	0.15	0.21	0.48
Baby Care Products	0.66	0.19	0.31	
Sanitary/Footcare Products	1.01	0.34	0.56	1.52
Perfume	0.35	0.88	0.27	

Table 2: CPI Category Weights (%)

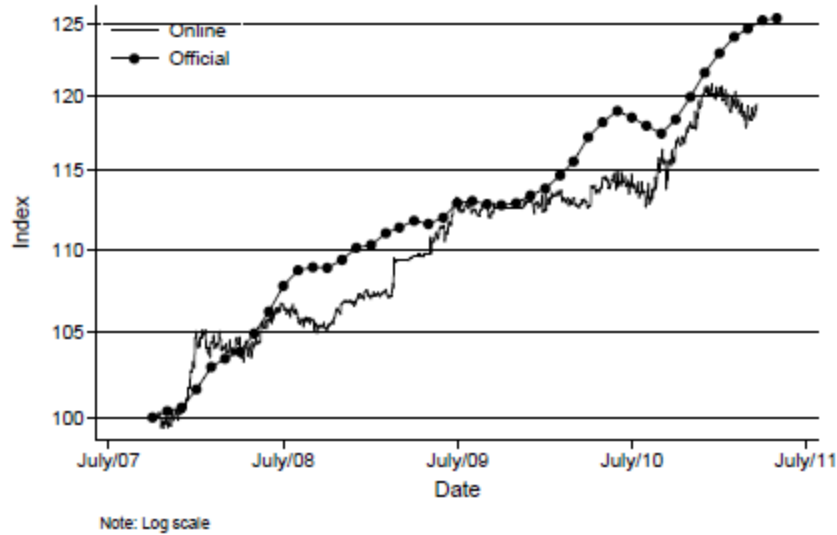
Category Name	Argentina	Brazil	Chile	Colombia
Flour And Prepared Flour Mixes	0.20	0.21	0.12	0.12
Cereals		0.35	0.19	0.19
Pasta	1.31	0.36	0.28	0.26
Rice	0.32	0.53	0.20	1.73
Bread	2.50	1.20	2.19	0.87
Fresh Biscuits, Rolls	1.31	0.08	0.04	
Cakes And Cupcakes	0.30	0.10	0.56	
Cookies	0.58		0.37	
Crackers And Bread & Cracker Products		0.47		0.58
Sweetrolls, Coffee Cake & Doughnuts	0.62	0.01		
Uncooked Ground Beef	0.29	0.01	0.04	
Uncooked Beef Steaks	5.76	1.50	1.89	2.48
Other Uncooked Beef And Veal		0.14		
Ham	0.54	0.12	0.60	
Pork Chops	0.11	0.17	0.34	0.50
Sausages	0.48	0.57	0.29	0.49
Lamb, Organ Meats, And Game	0.14		0.03	
Chicken	2.18	0.90	0.89	1.31
Other Poultry Including Turkey			0.07	
Appliances	1.05	1.21	1.38	0.51
Books	0.80	0.30	0.13	0.07
Home Furniture	0.73	1.33	1.44	0.46
Bedroom And Bathroom Linens		0.27	0.32	
TOTAL WEIGHT (% of CPI)	48.51	27.93	31.00	28.44

Methodology

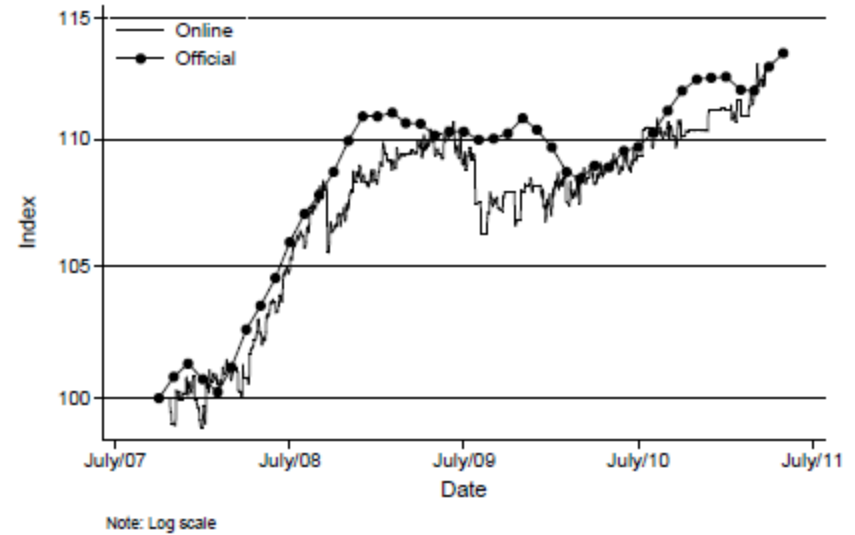
In addition to the index, I provide estimates of:

- Monthly inflation:
% change in the average of the index for the past 30 days vs the previous 30 days
- Annual inflation
% change in the average of the index for the past 30 days vs the average for the same period a year ago

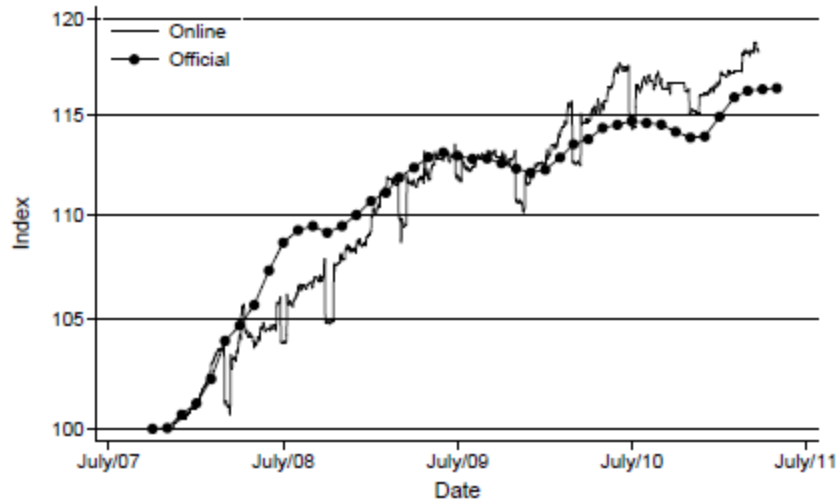
(a) Brazil



(b) Chile



(c) Colombia



(d) Venezuela

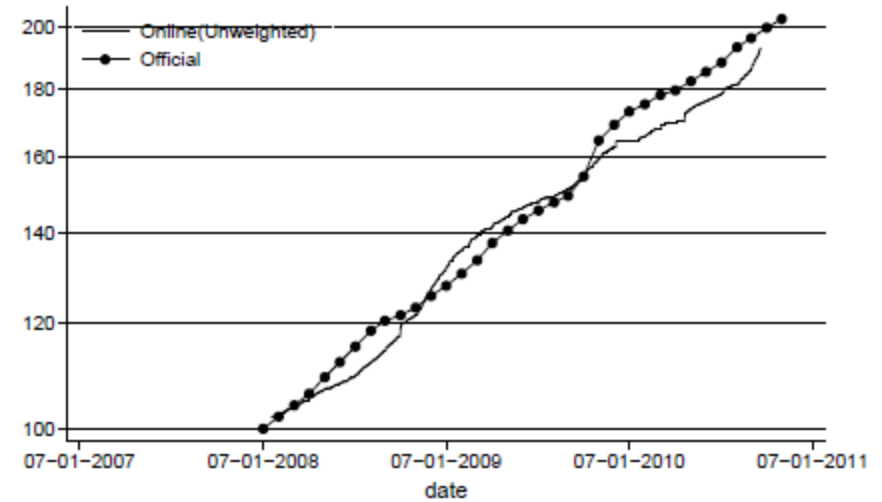
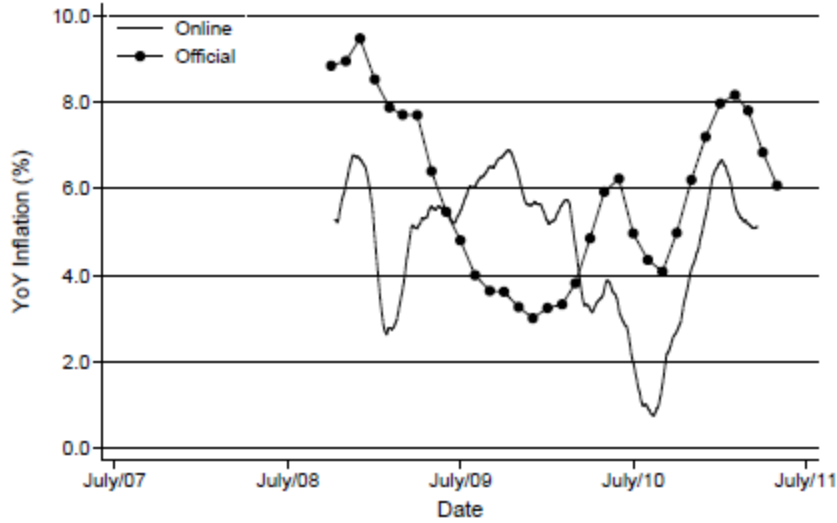
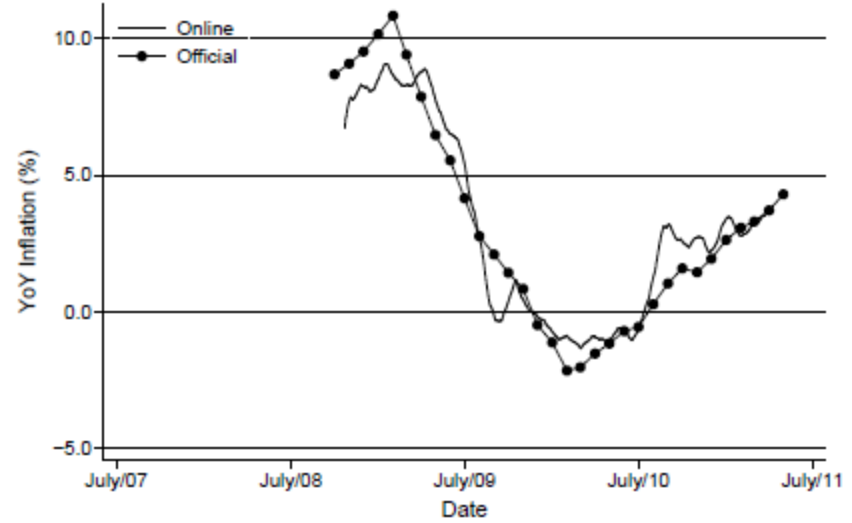


Figure 1: Online and Official Indexes in Four Latin American Countries

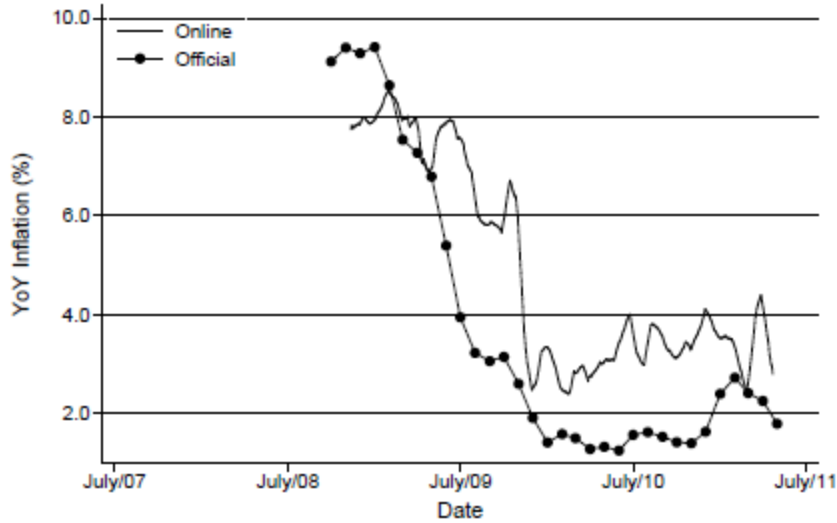
(a) Brazil



(b) Chile



(c) Colombia



(d) Venezuela

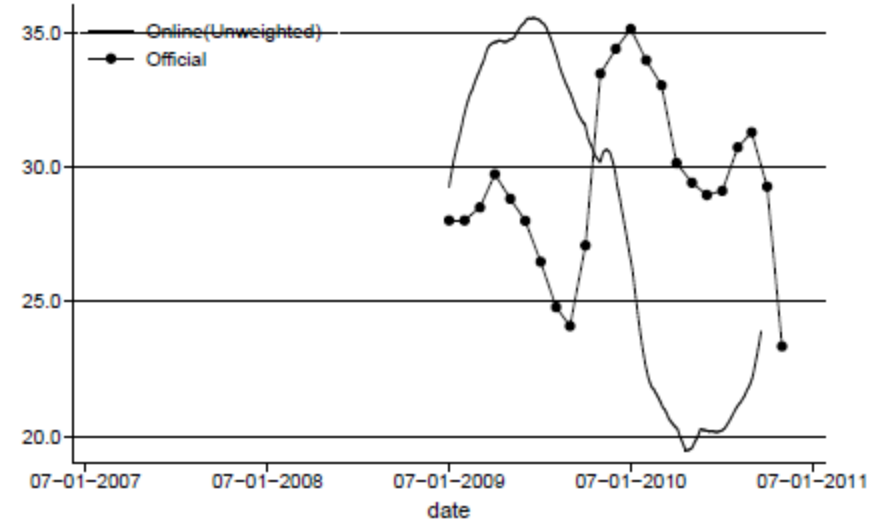
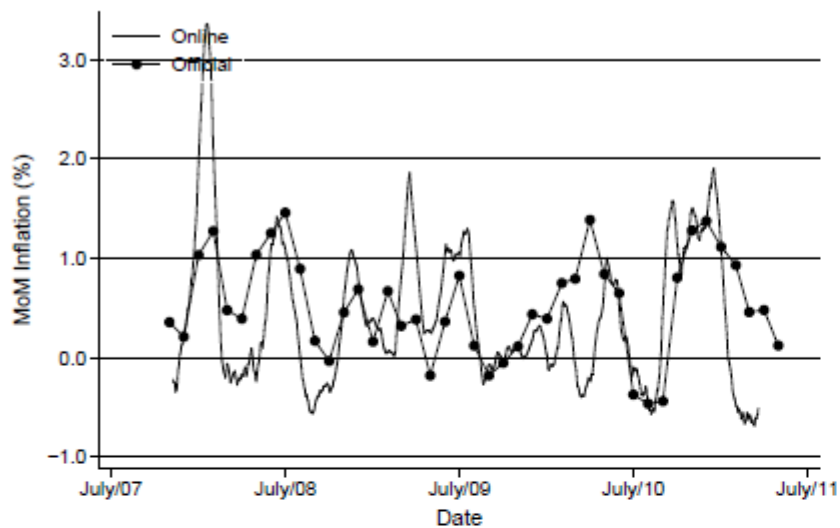
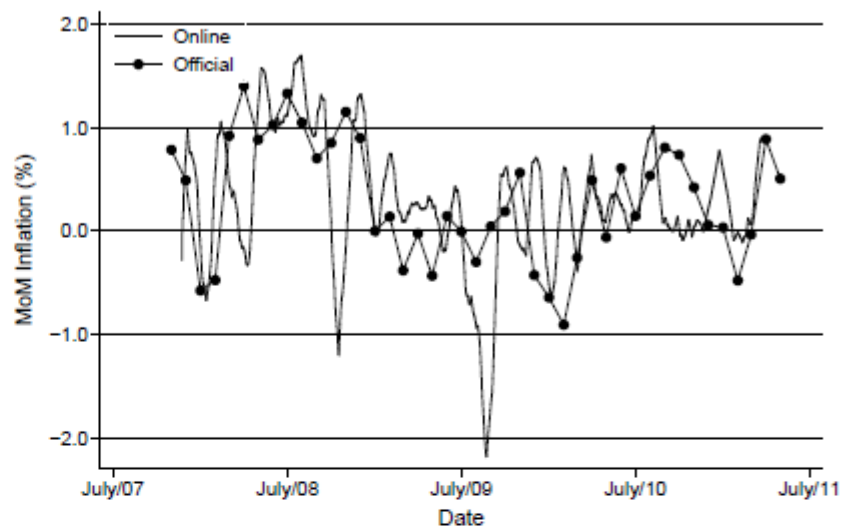


Figure 2: Online and Official Indexes - Annual Inflation Rate

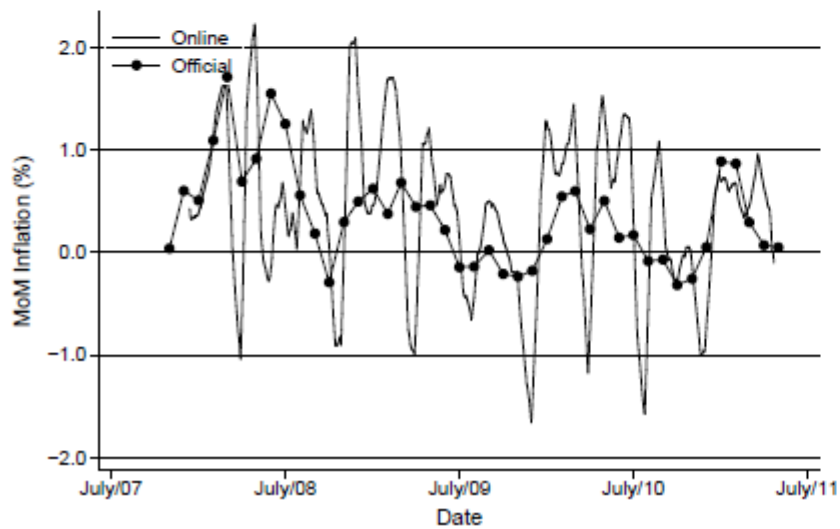
(a) Brazil



(b) Chile



(c) Colombia



(d) Venezuela

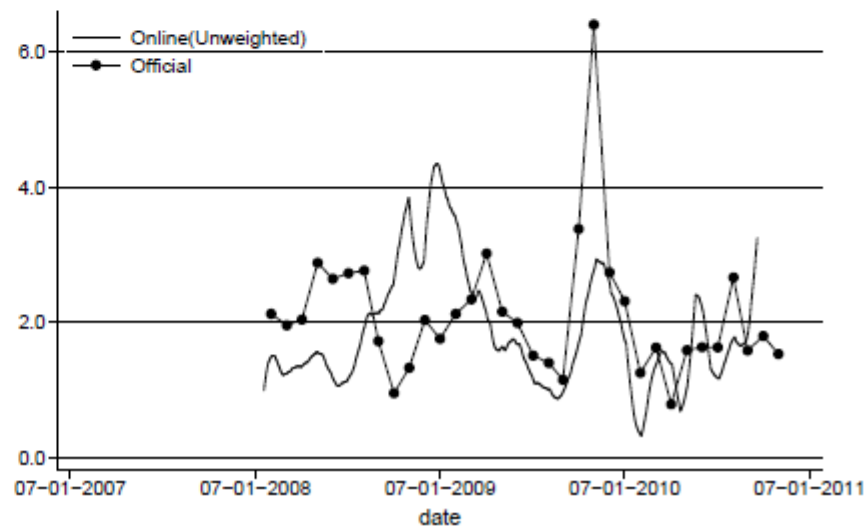


Figure 3: Online and Official Indexes - Monthly Inflation Rate

Table 3: Online vs Official Series

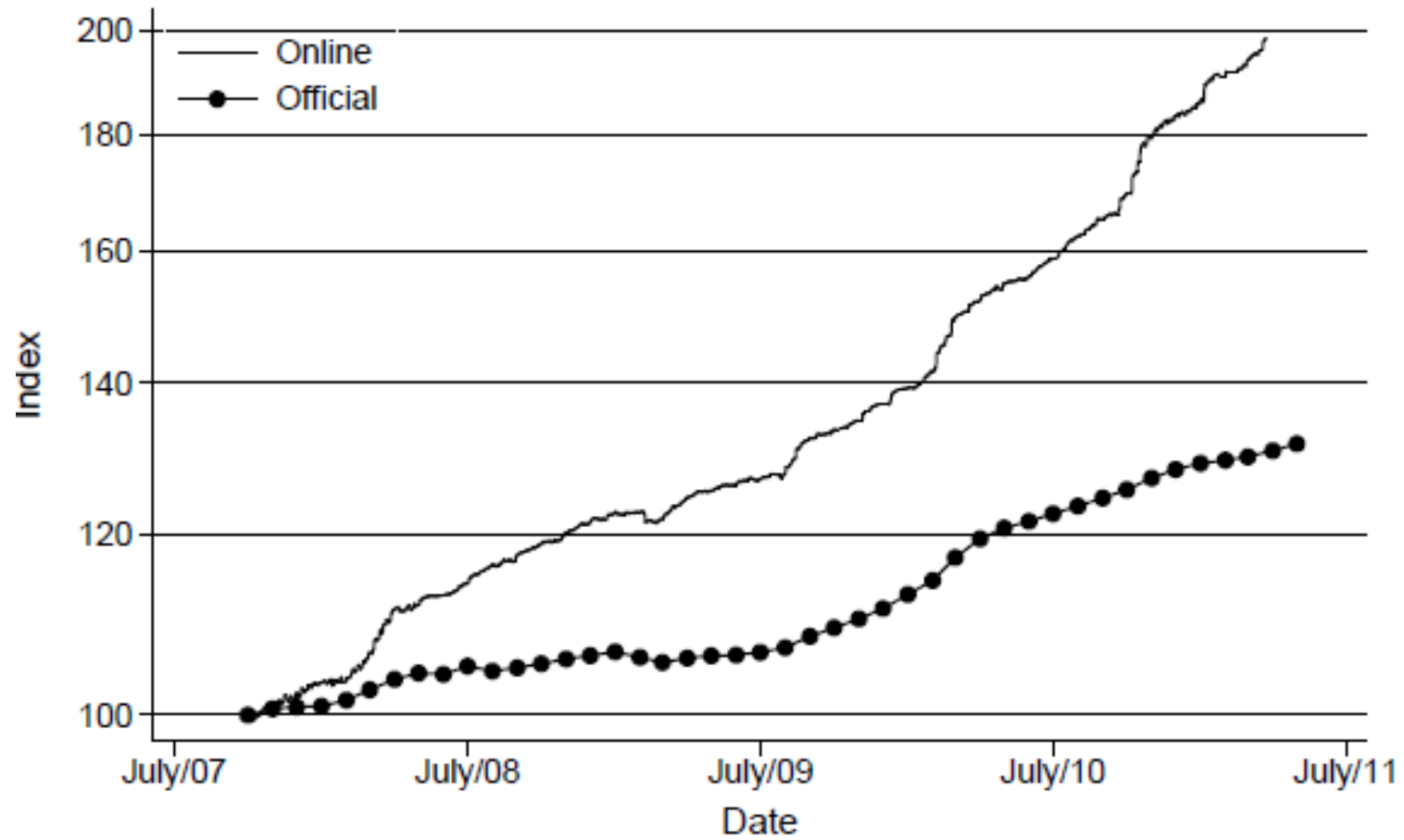
	Argentina	Brazil	Chile	Colombia	Venezuela
Mean Annual Inflation (%)					
Online Index		4.72	3	4.88	27.43
Official Index		5.91	3.19	3.73	29.38
Contemporaneous Correlation Online and Official					
Price Index		0.96	0.97	0.95	0.92
Monthly Inflation		0.5	0.38	0.43	0.18
Annual Inflation		0.09	0.97	0.89	-0.08

Brazil, Chile, Colombia, Venezuela

- Online indexes are able to track main inflation trends
 - Even with only one retailer in each country
- Matching is best In Chile and Colombia, where:
 - Supermarkets have larger market shares (27% and 30%, vs. only 15% in Brazil)
 - City concentrates population & accounts for most of the CPI (55% in Santiago, Chile)
- → Good news for Argentina!...the supermarket used has 28% market share, Buenos Aires is 100% of CPI data

Argentina

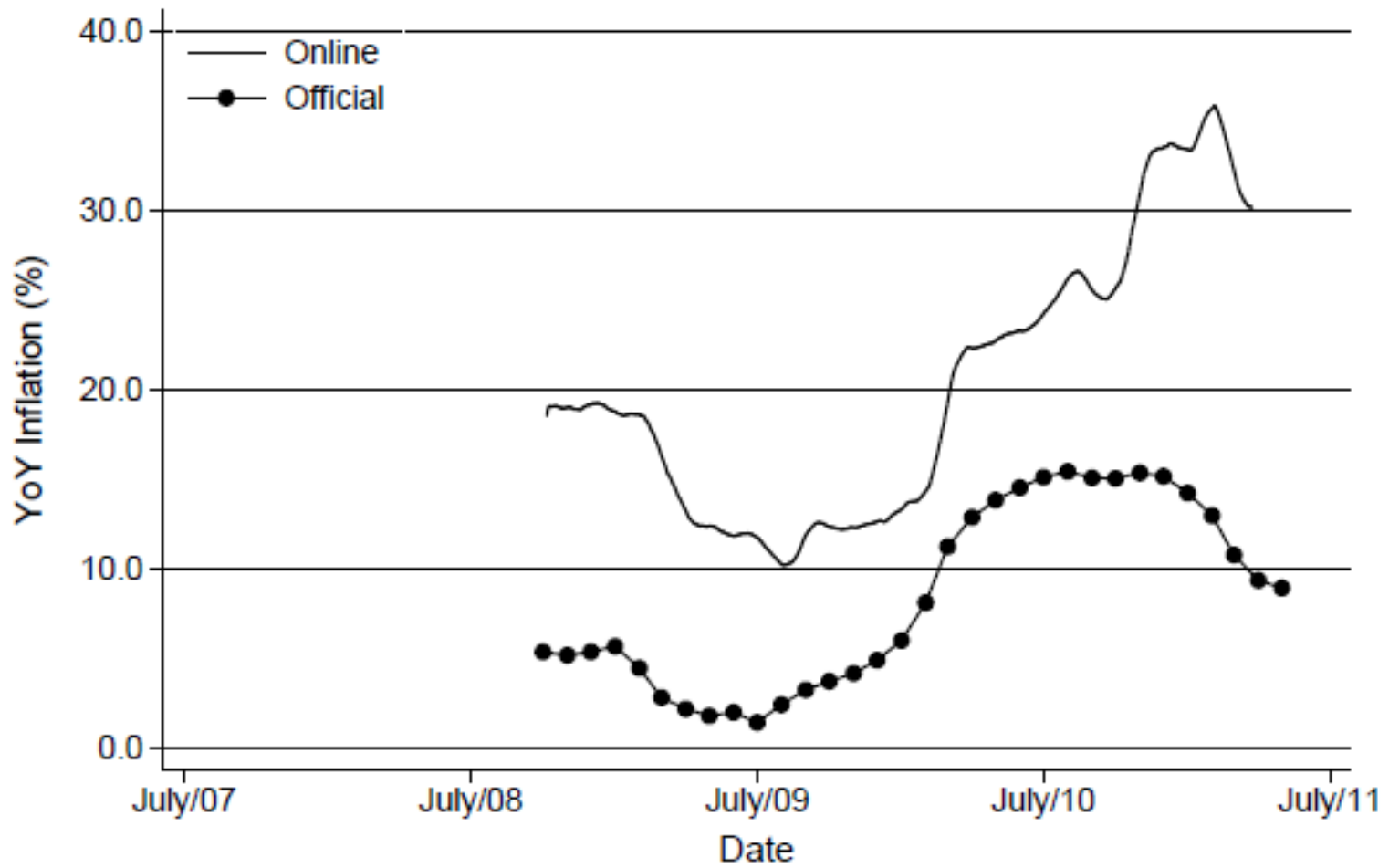
(a) Daily Index



Note: Log scale

Argentina

(b) Annual Inflation



Argentina

(c) Monthly Inflation

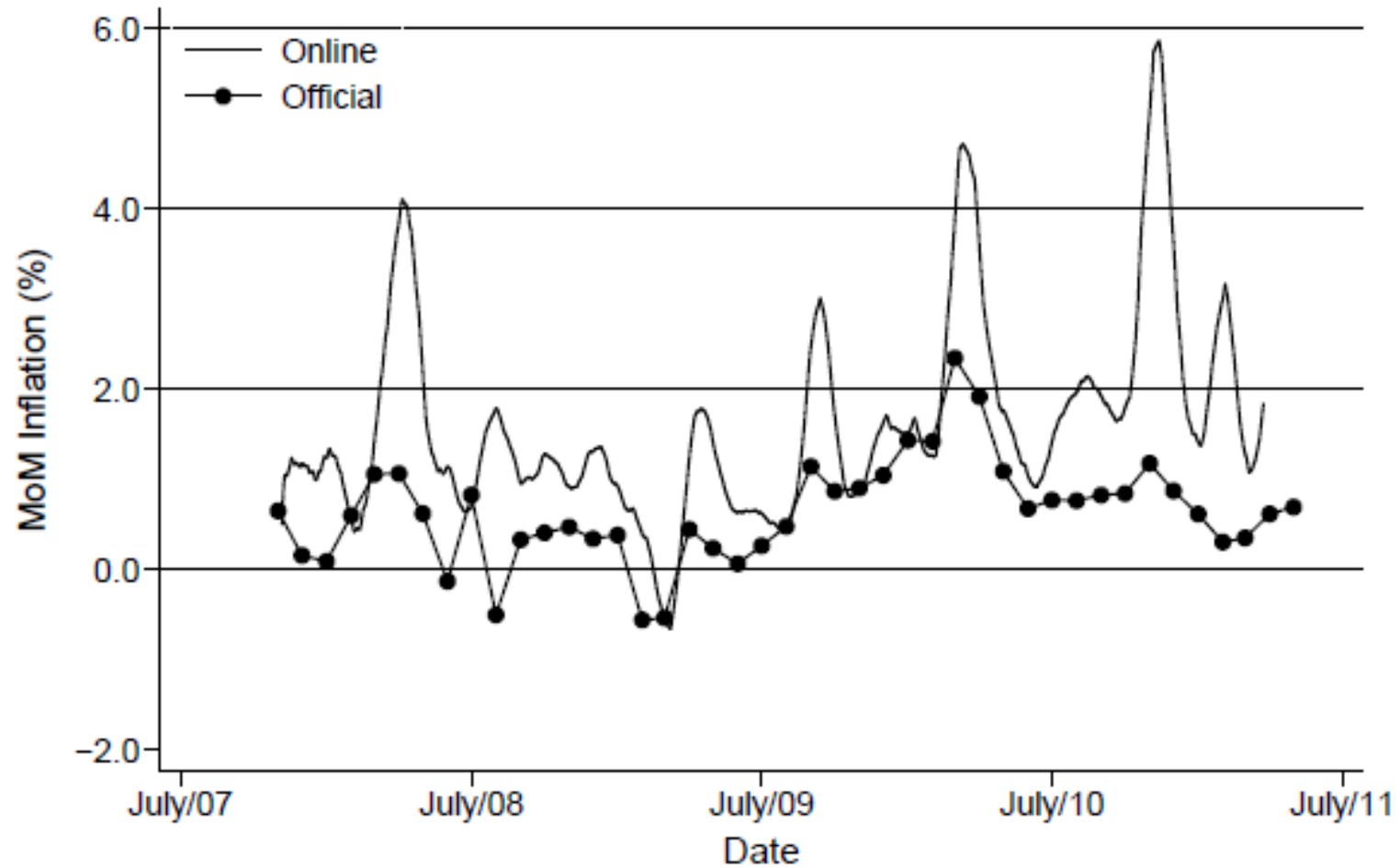


Table 2: Online vs Official Series

	Argentina	Brazil	Chile	Colombia	Venezuela
Mean Annual Inflation (%)					
Official CPI Index	8.53	5.28	2.44	3.79	27.37
Official Supermarket Index	8.38	5.91	3.19	3.73	29.38
Online Supermarket Index	20.14	4.72	3	4.88	27.43
Correlations between Online and Official Supermarket Series					
Price Index	0.98	0.96	0.97	0.95	0.92
Annual Inflation	0.84	0.09	0.97	0.89	-0.08
Monthly Inflation	0.6	0.5	0.38	0.43	0.18
Regression - Official Supermarket on Online Monthly Inflation Rates (12 lags)					
Constant	0.84	-0.54	0.14	0.03	-1.96
Constant p-value	0.000	0.19	0.17	0.86	0.23
R2	0.9	0.55	0.66	0.59	0.66
Monthly Inflation Rate Volatility (standard deviation)					
Official Supermarket Index	0.57	0.51	0.58	0.48	0.98
Online Supermarket Index	1.11	0.73	0.62	0.76	0.86

What can explain the differences?

- Scraped data or the index methodology?
 - Similar scraped-data indexes can closely track official CPI statistics in 4 other Latin American countries.
 - Same for Subsistence Food Index (CBA) → simpler methodology, carefully selected items

Subsistence Food Index - CBA

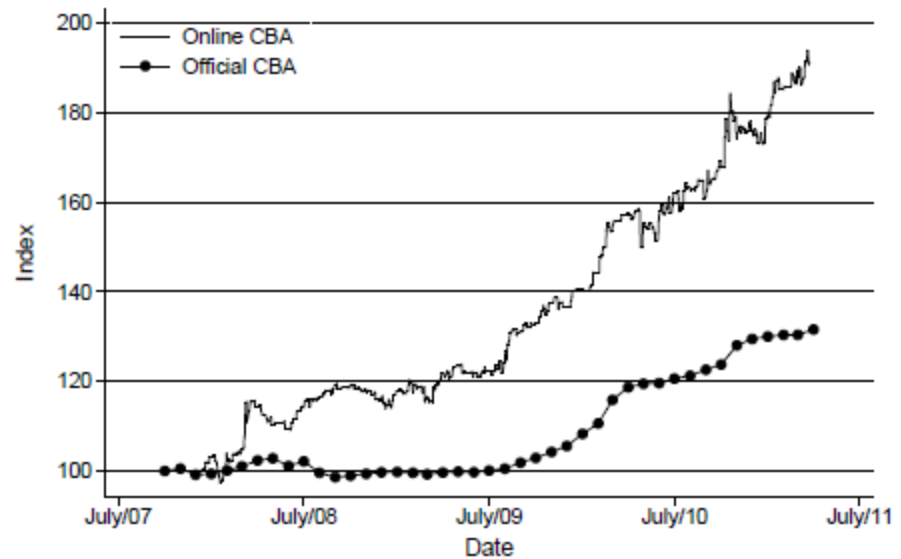
- Used by INDEC to calculate level of extreme poverty
- Methodology
 - I replicate INDEC's methods in 2007
 - Basket of basic food → minimum nutrition requirements adult male
 - Compute cost each day → 45 goods, using \$/gram
 - Simple Laspeyres Index $I_t = \text{cost at } t / \text{cost at } t_0$

Subsistence Food Index (CBA)

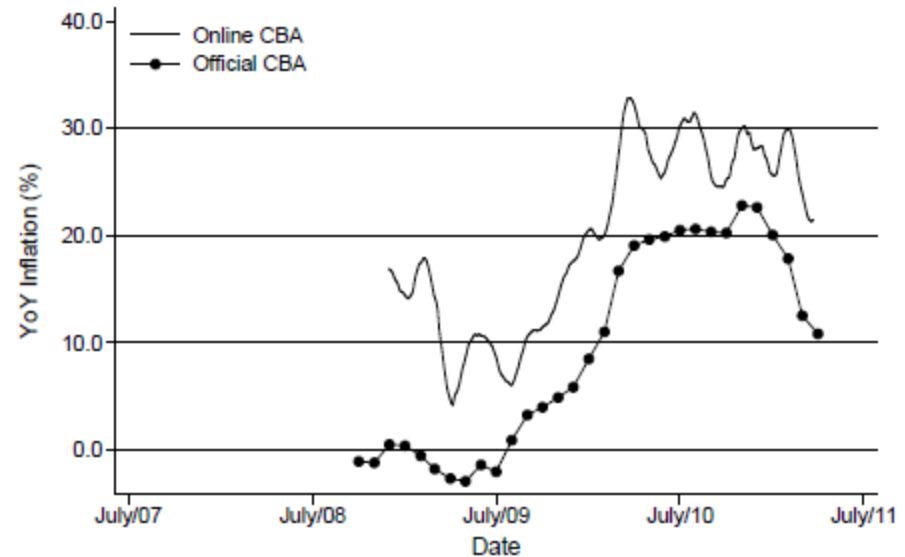
Implications for poverty: take the cost of the CBA basket in early 2008 and adjust it with the inflation rate observed online:

- Extreme poverty is 6.69% instead of 2.5% officially reported
- Poverty is 25.9% instead of 9.9% officially reported

(a) Daily Index



(b) Annual Inflation



Subsistence Food Index - CBA

- Implications for poverty estimates: if we take the cost of the CBA basket in early 2008 and adjust it with the inflation rate observed online:
 - Extreme poverty is 6.69% instead of 2.5% officially reported
 - Poverty is 25.9% instead of 9.9% officially reported

What can explain the differences?

- Online prices have more inflation?
 - Not in other countries
 - 2-3x more inflation for 3.5 years?
 - Survey of online and offline prices for this retailer found constant markup of 5% in online prices → implies same inflation rates

Table 8: Online vs. Offline Prices

	Argentina	Brazil	Chile	Colombia
Matching ids	Yes	Yes	No	Yes
% Available Online	100%	80%	90%	74%
PRICE LEVELS				
online=offline	18%	42%	93%	29%
online>offline	78%	34%	4%	32%
Price Difference (Mean %)	5	9	2	0
PRICE CHANGES				
Products with Identical Change Series*	93%	75%	94%	67%
Ratio of Changes over Observations				
Offline	0.215	0.356	0.274	0.433
Online	0.215	0.411	0.249	0.433
Mean Size of Changes (%)				
Offline	1.6	4.9	1.4	8.1
Online	1.4	5.3	1.3	8.3

Notes: *Indicator variable conditional on change: 1 if the price increased, -1 if it decreased.

Taken from Cavallo (2010) "Scraped data and Sticky Prices"

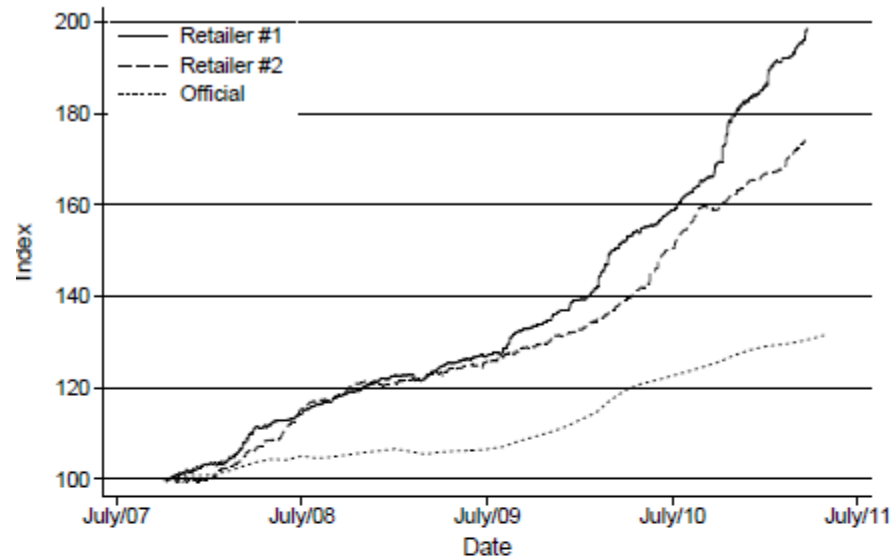
What can explain the differences?

- Is this supermarket “special”?
 - Similar results in two supermarkets with very different characteristics

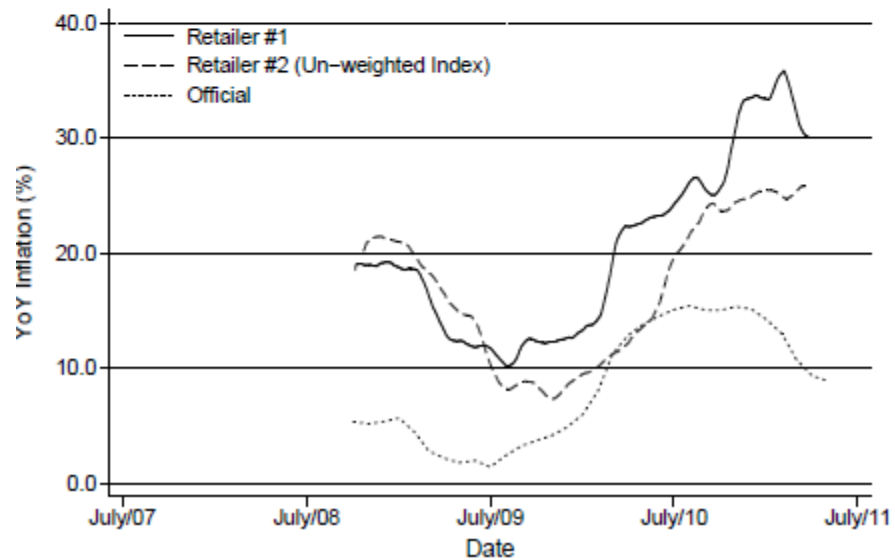
Two Supermarkets

Supermarket #2 is small, only sells online, and caters to high-income consumers

(a) Daily Index



(b) Annual Inflation



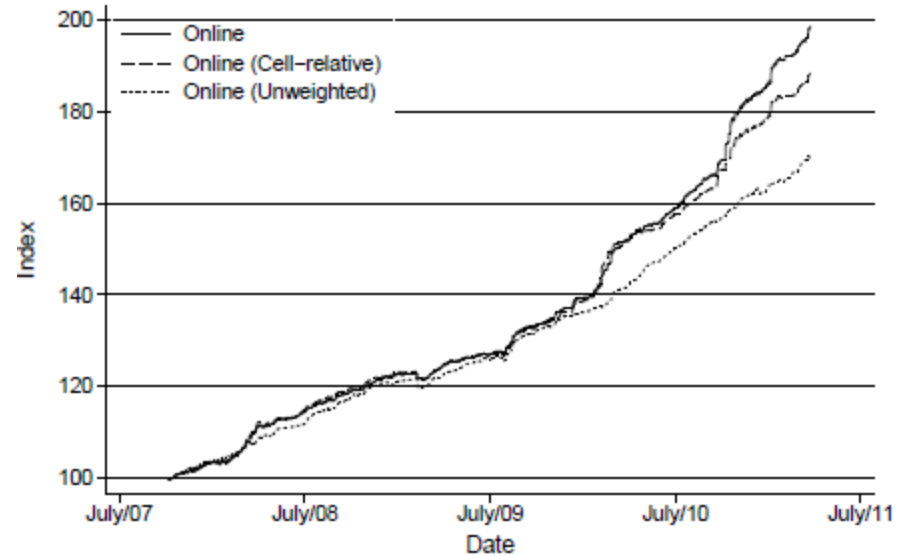
What can explain the differences?

- Other robustness
 - Cell-relative imputation of missing values (BLS-style)
 - Un-weighted index
 - Price-controlled goods
 - Low-inflation goods

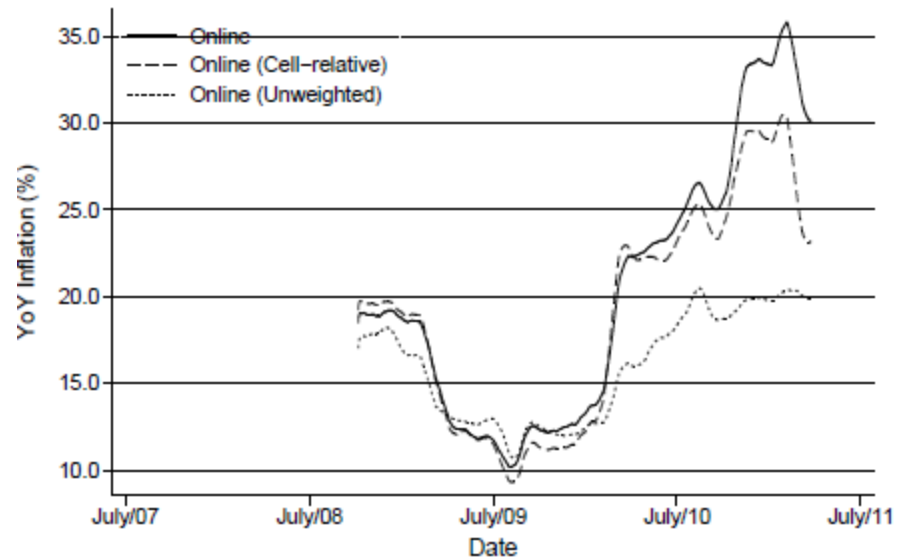
Other Methods

- Cell-relative imputation of missing values (BLS-style)
- Un-weighted index

(a) Daily Index



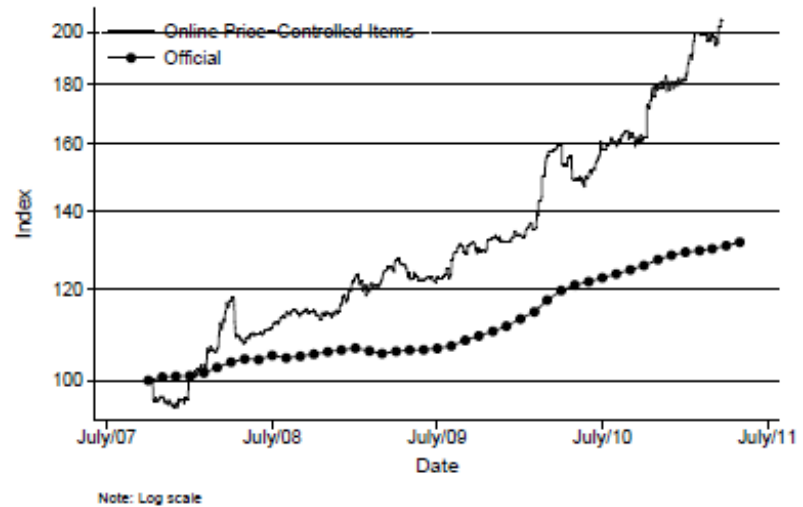
(b) Annual Inflation



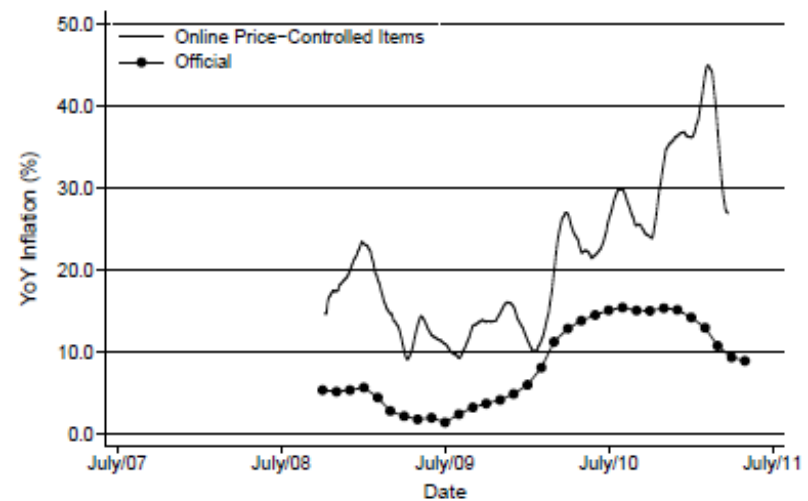
Price-Controlled Products

- Using only the products that were under price-controls at some point
- 597 products

(a) Daily Index



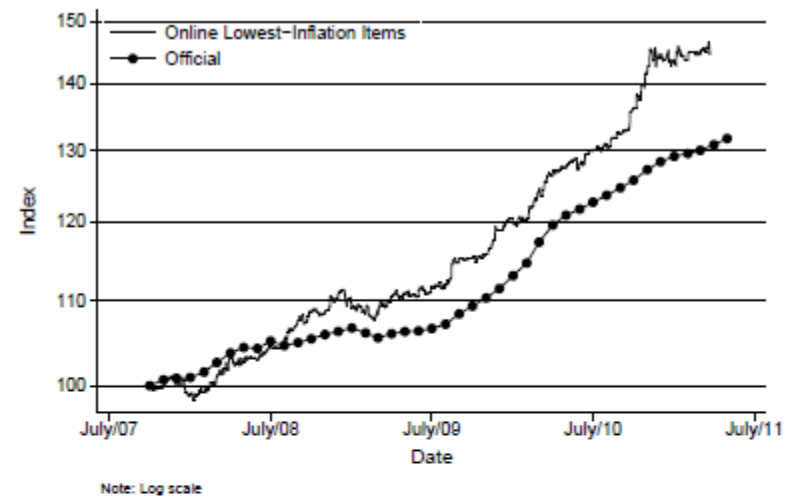
(b) Annual Inflation



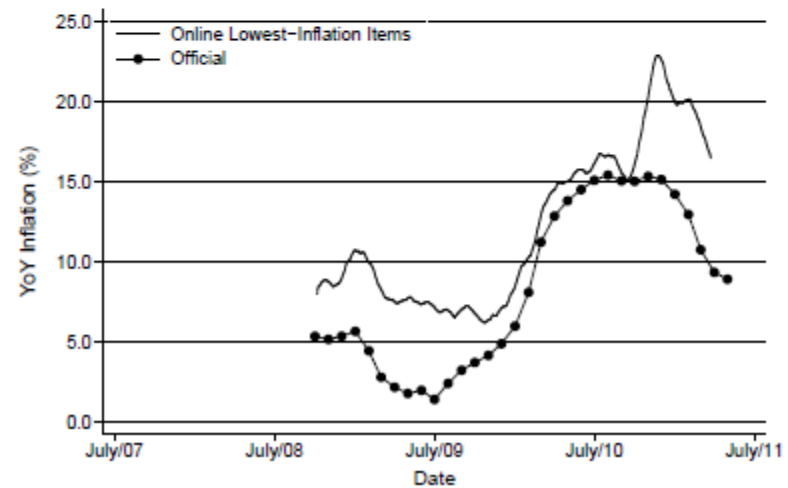
Lowest – Inflation Products

- Using the products with the lowest inflation for the whole period within each URL (ex-post)
- 893 products

(a) Daily Index



(b) Annual Inflation



Benford's Law

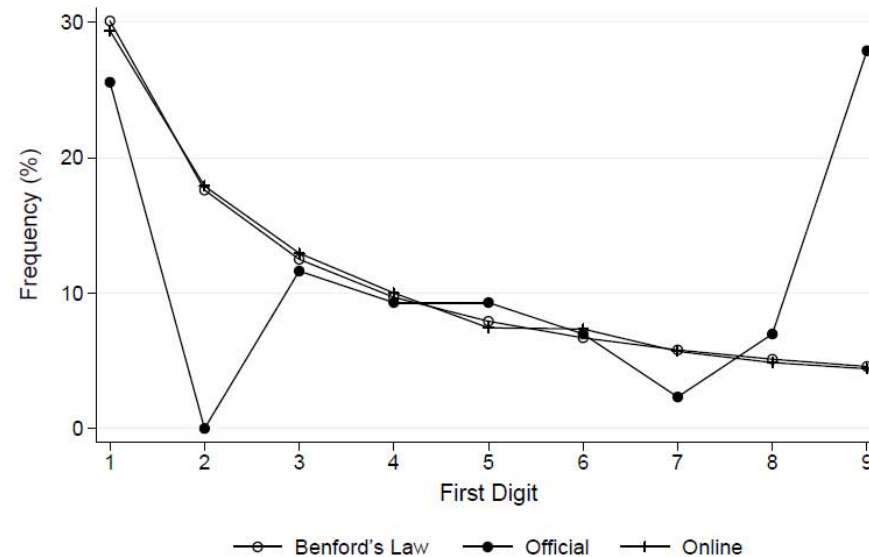


Figure A1: Benford's Law

Notes: Benford's Law is an empirical regularity commonly used to detect accounting fraud. It stipulates that the first digit in a list of numbers describing naturally-occurring phenomena should follow a specific non-uniform distribution (eg. 30.1% for 1, 16.6% for 2, etc.). The figure plots the results of calculating the frequency of each first digit in the series given by the change in online and official price indexes. Online data appears to satisfy Benford's Law, but not the official series, which have an excess of 9's and lack of 2's.

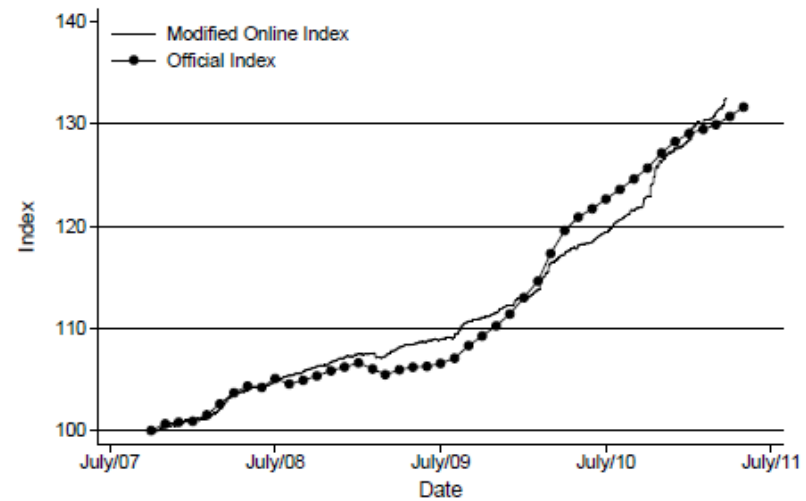
Best Approximation

- Assuming there is manipulation, how is this taking place?
 - At the good-level, when prices are collected?
 - No product-level prices published, but this is unlikely, because it requires the cooperation of people doing the price collection
 - At the aggregate-level, when the final index is computed?
 - On average, online inflation is $1/2$ to $1/3^{\text{rd}}$ of official data, so let's try dividing the observed daily inflation by 3

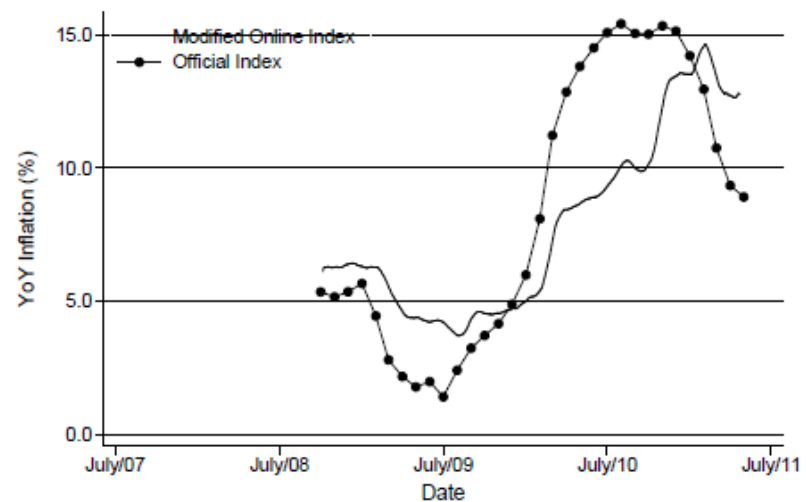
Best Approximation

- Use 1/3rd of the online inflation rate

(a) Daily Index

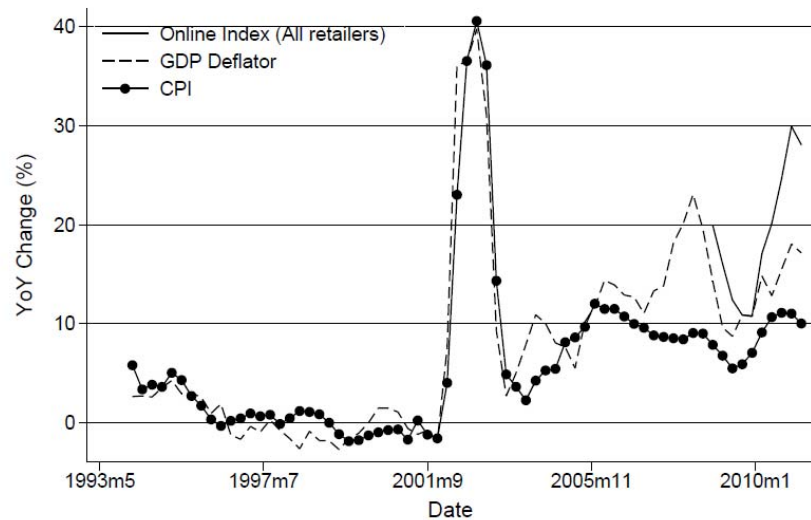


(b) Annual Inflation

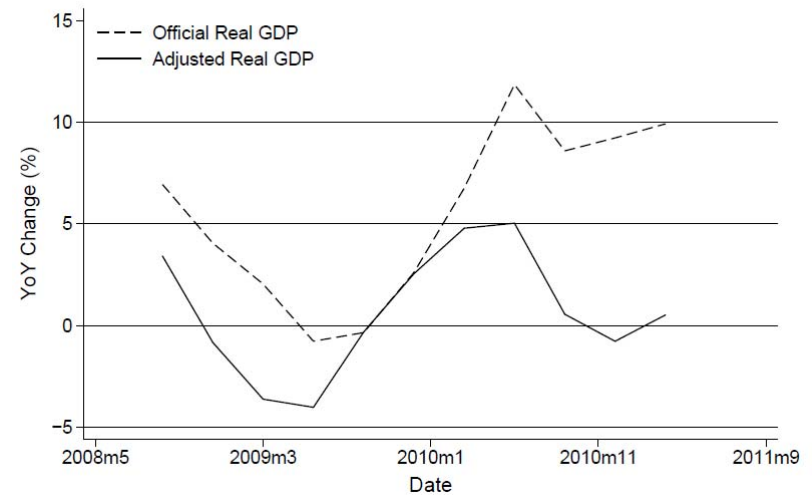


Implications for Growth Estimates?

(a) GDP Deflator, CPI, and Online Index - Annual Change



(b) Real GDP - Annual Growth Rate



Conclusions

- Inflation in Argentina is between 2 or 3 times higher than officially reported.
- Differences are not due to data or methodology.
- Argentina's manipulation of official indexes appears to be the exception across countries
- Scraped data → useful to construct daily price indexes that are updated on a real-time basis, are faster, cheaper, and can complement official statistics

More on Online Price Indexes

- Inflation Verdadera

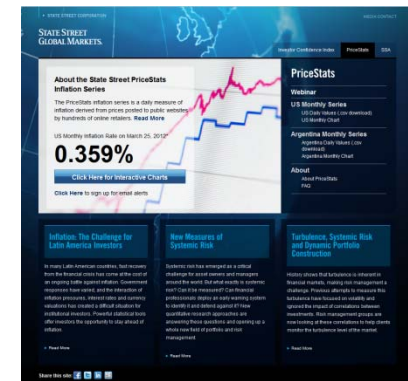
- www.inflacionverdadera.com



- Argentina's PriceStats Index

- Published in The Economist every week

- <http://www.statestreet.com/research/pricestats/index.html>



- Billion Prices Project at MIT

- bpp.mit.edu

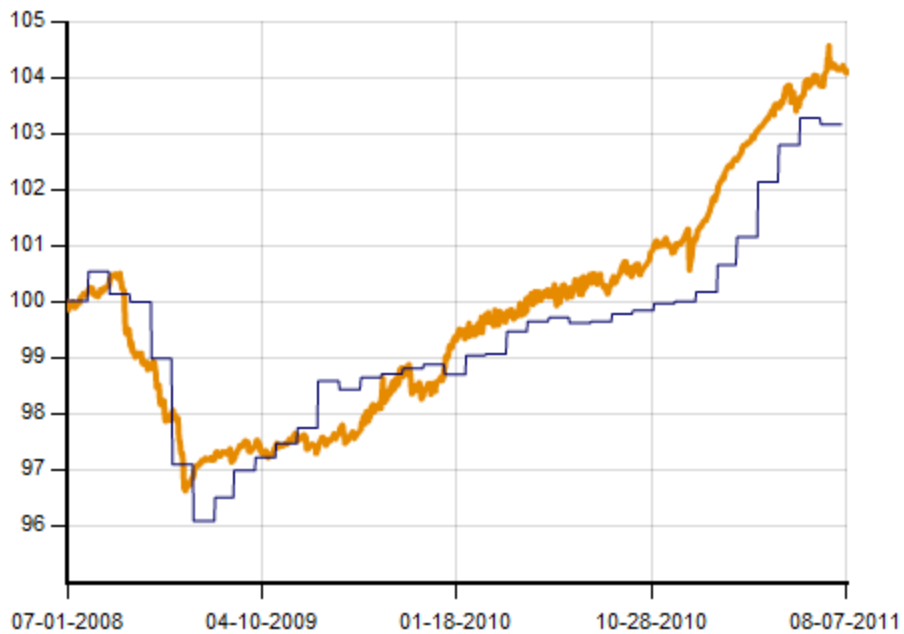


Alberto Cavallo
MIT Sloan

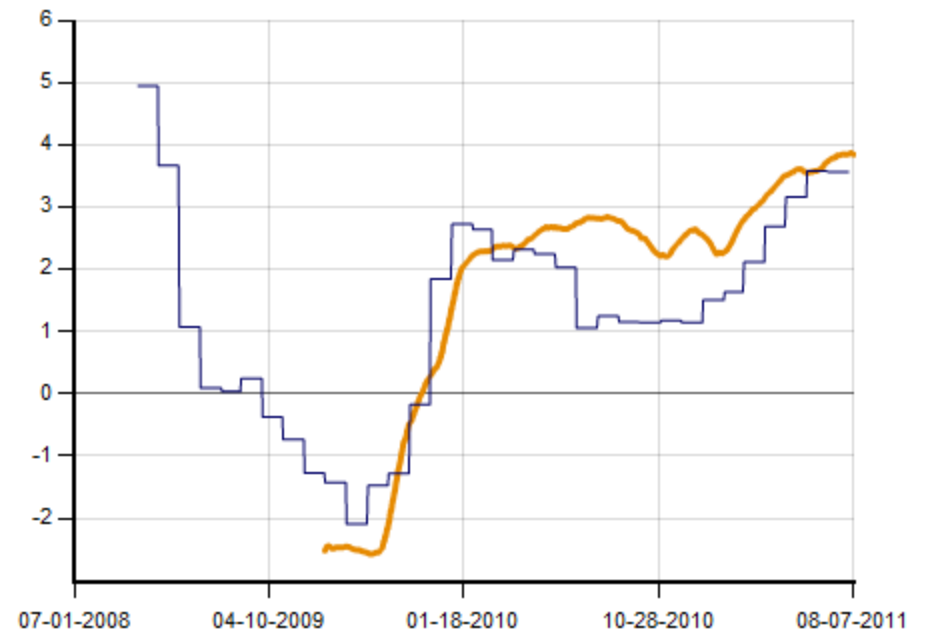
acavallo@mit.edu
acavallo.mit.edu
bpp.mit.edu

US Index

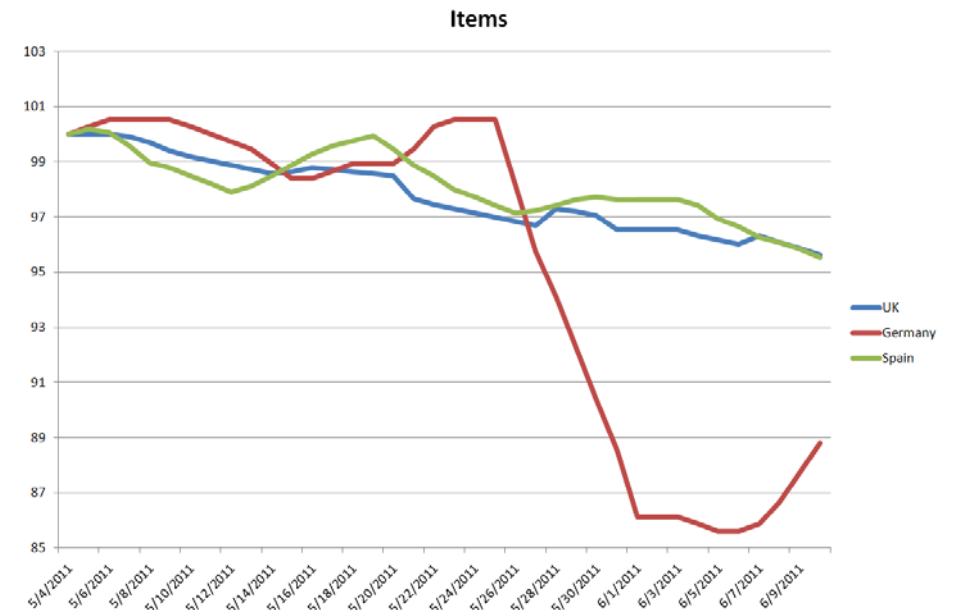
PRICE INDEX



ANNUAL INFLATION



E-coli Crisis in Europe



Effects of Japan's Earthquake

