

Real-Time Inflation Measurement

Alberto Cavallo
Harvard Business School

CFM - PER Data Initiative
Columbia University March 2021

Big Data in Macro and International Economics

- Macro data have many problems (Eichengreen (2015) , Summers (2016))
- We mostly rely on data collected by governments → Griliches (1985) The Uneasy Alliance

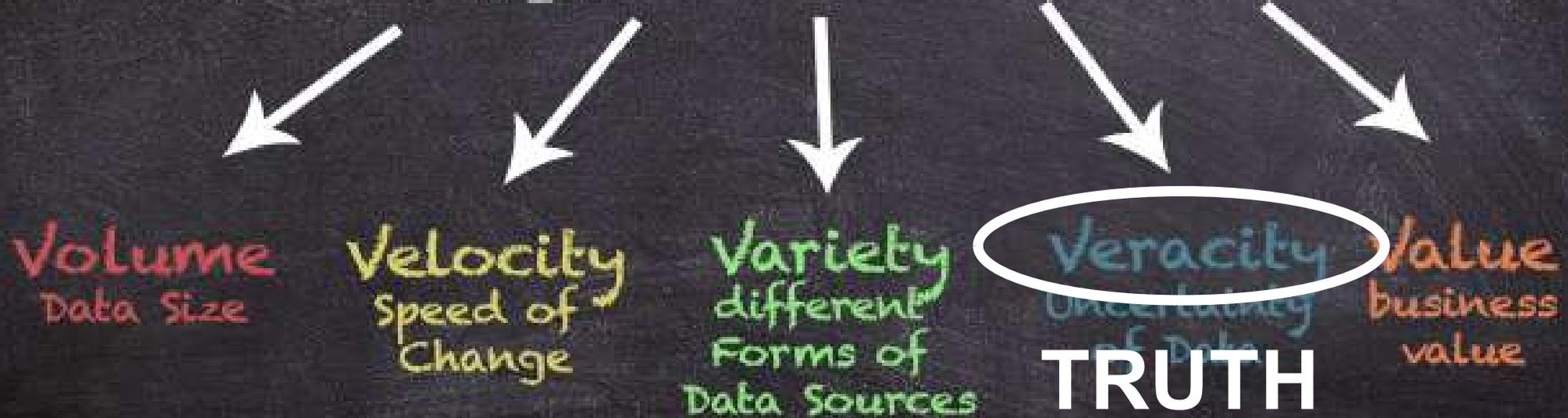
“... we have shown little interest in improving it [the data], in getting involved in the grubby task of designing and collecting original data sets of our own. Most of our work is on “found” data, data that have been collected by somebody else, often for quite different purposes... “They” collect the data and are responsible for all their imperfections. “We” try to do the best with what we get, to find the grain of relevant information in all the chaff.”

- Can Big Data help? (Einav & Levin (2014))

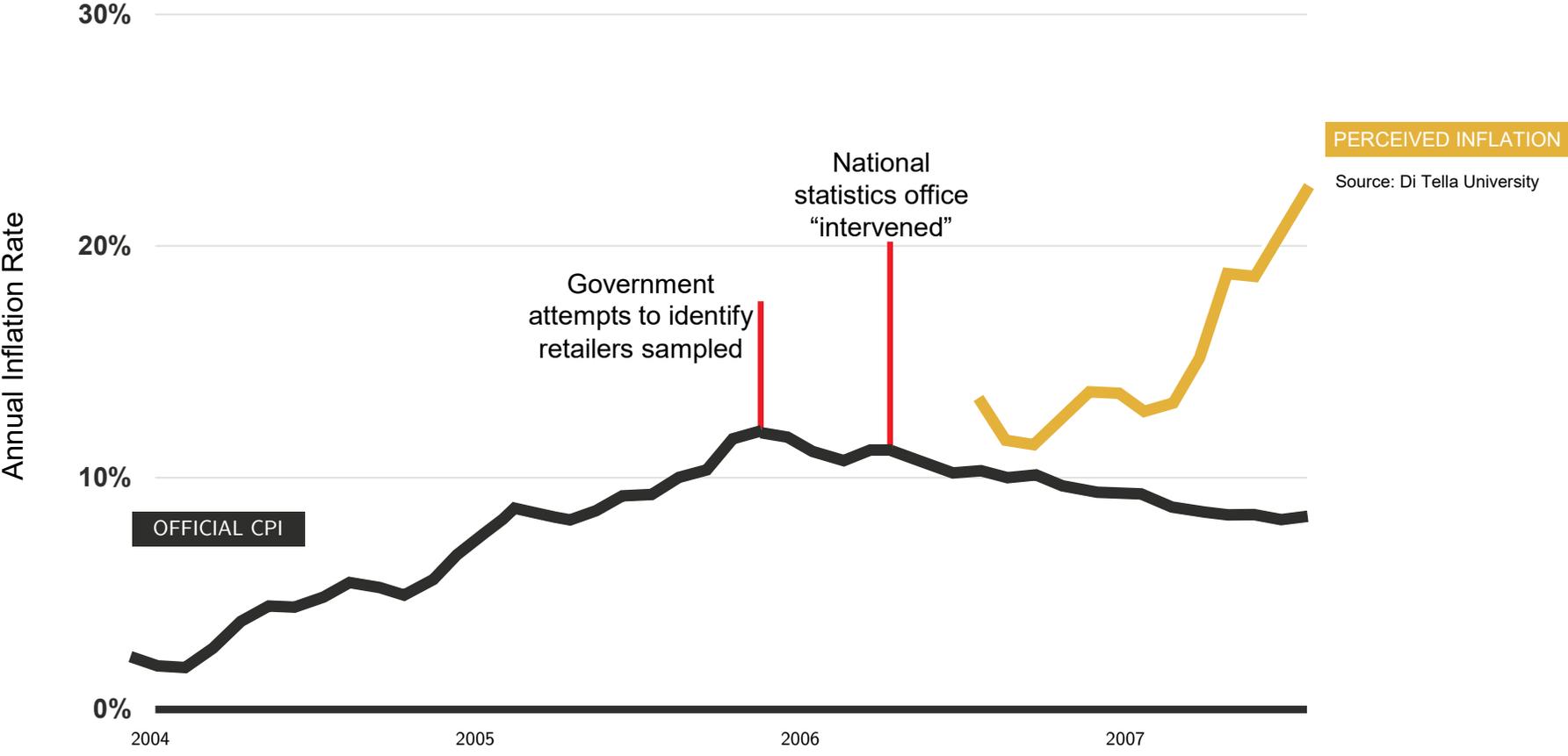
Big Data in Macro and International Economics

- New data sources in recent years
 - Administrative data (eg. CPI and IPI micro data, tax & property records)
 - Scanner data (eg. Nielsen)
 - Search data (eg. Google, Indeed)
 - Satellite data (eg. lights, parking lots, tanker and crop heights)
 - Sensor data (smart phones, smart watches, IOT devices)
 - Crowd-sourced data (web, mobile phones)
 - Online data (eg. Billion Prices Project)
- The 5 Vs and the origin of the BPP

Big Data



DECEPTION IN ARGENTINA, 2007-2015



Source: Di Tella University

ONLINE PRICES PROVIDED AN ALTERNATIVE WAY TO COLLECT PRICE DATA

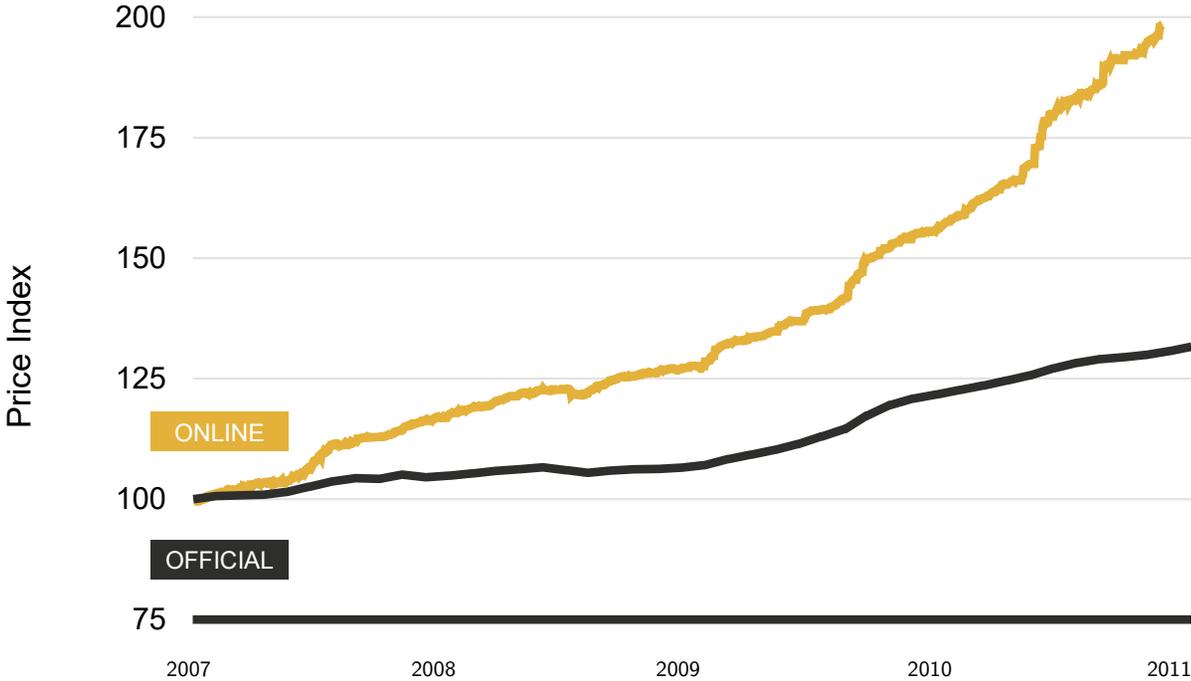
Producto	Descripción	Precio	Cantidad	Comprar
	Leche Condensada Nestlé Pack 3 unidades, Lata 200 grs. c/u SiLitro: \$1.198	\$1.199 Uni	<input type="text"/> Uni	
	Leche Evaporada Ideal Lata 400 grs. SiLitro: \$1.473	\$099 Uni	<input type="text"/> Uni	
	Leche Evaporada Jumbo Lata 410 grs. SiLitro: \$1.193	\$099 Uni	<input type="text"/> Uni	
	Leche Condensada Nestlé Envase flexible 350 grs. SiLitro: \$1.549	\$099 Uni	<input type="text"/> Uni	
	Leche Condensada Nestlé Descremada, Lata 395 grs. SiLitro: \$1.983	\$799 Uni	<input type="text"/> Uni	

→ `<html>`
`<descripcion> Leche Condensada </descripcion>`
`<brand> Nestlé </brand>`
`<td price> $1.199 Uni </td>`

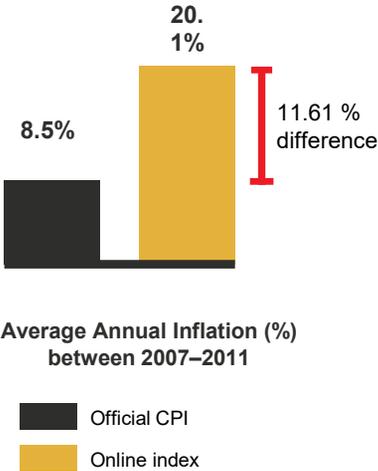
↓

	ID	ID2	PRODUCT	BRAND	SIZE	BULK PRICE	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	xKilo:\$1.998	1.199

ONLINE INFLATION IN ARGENTINA



Source: Cavallo (2012) "Online and Official Price Indexes: Measuring Argentina's Inflation", Journal of Monetary Economics



Online indices matched CPIs in other countries

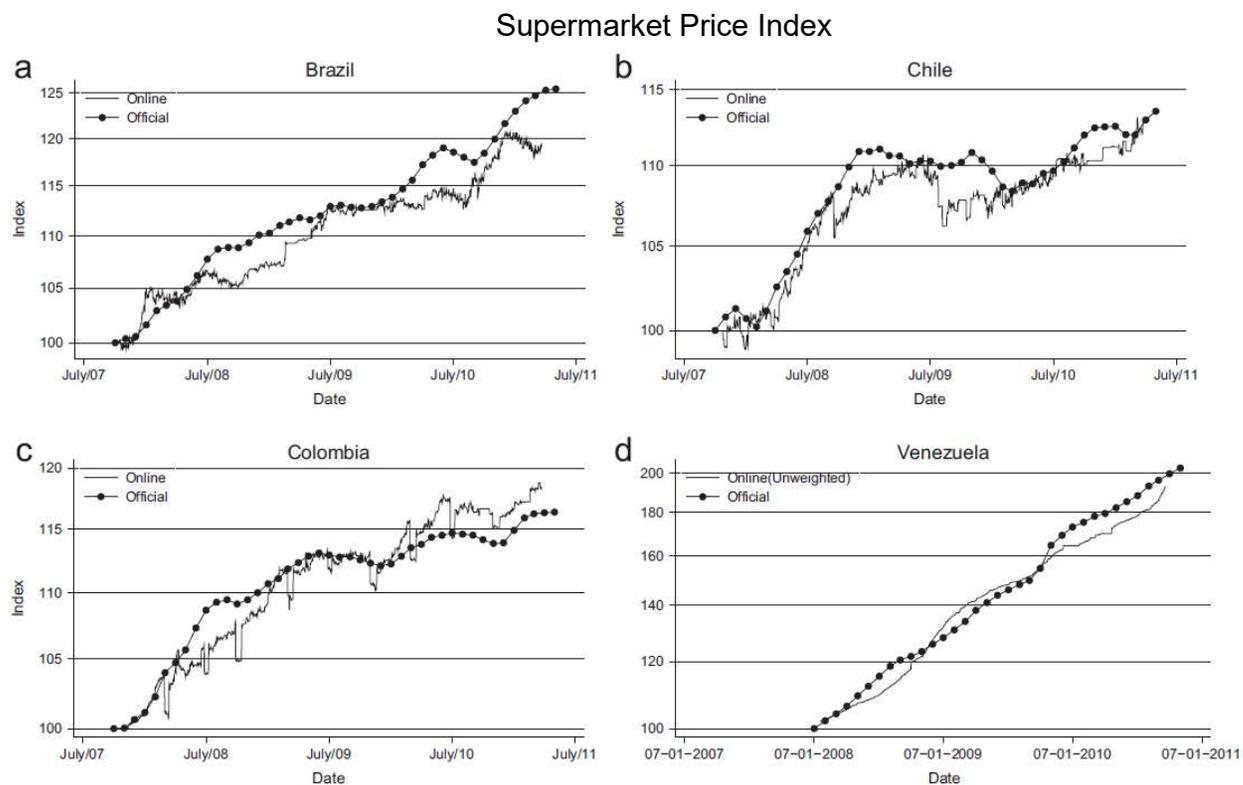


Fig. 1. Online and official indexes in four Latin American countries: (a) Brazil; (b) Chile; (c) Colombia; and (d) Venezuela. *Notes:* The daily online supermarket index is constructed with an online prices and official CPI category weights. In Venezuela, the online data has no category information and therefore the online index is built as a geometric average of all price changes observed each day. The official supermarket index is an equivalent indicator constructed with official CPI category weights. *Source:* Cavallo (2013) Online vs. Official Price Indexes: Measuring Argentina's Inflation. *Journal of Monetary Economics*, Vol. 60, No. 1, pp. 1-15.

IS IT ONLINE DATA?



Venezuela



Colombia



Chile



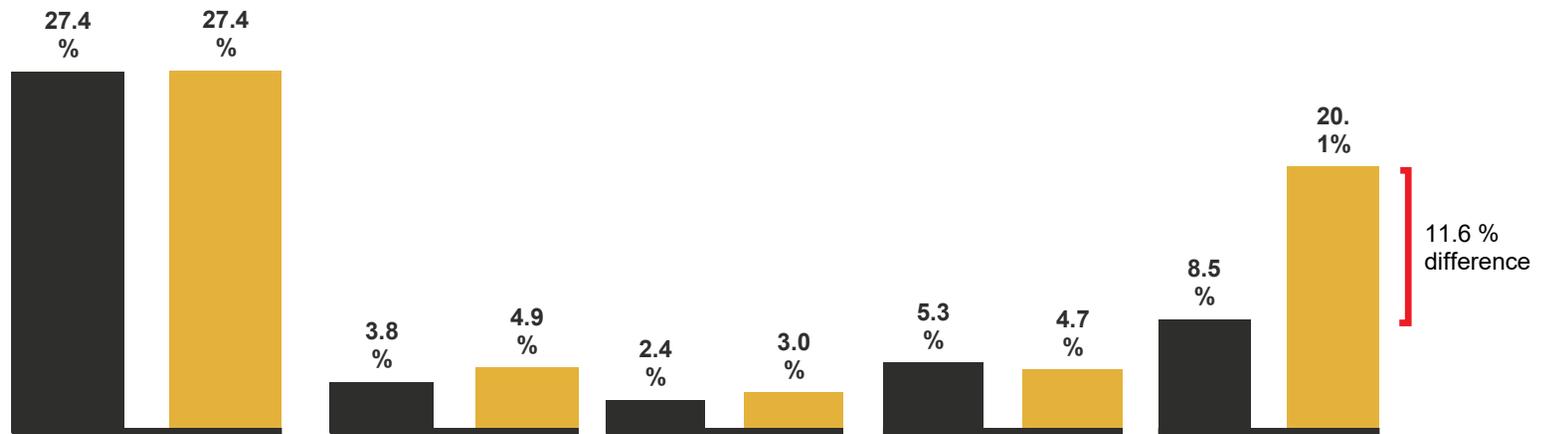
Brazil



Argentina

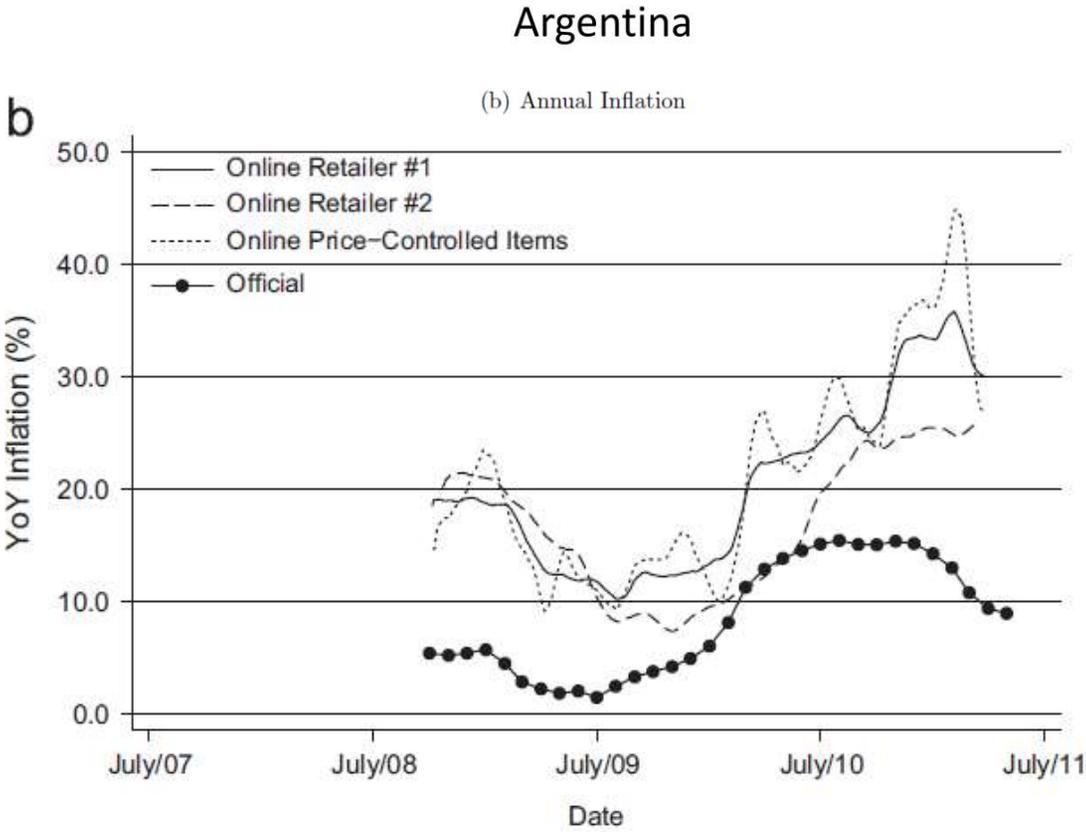
Average annual inflation (%)
between 2007–2011

Official CPI
Online index



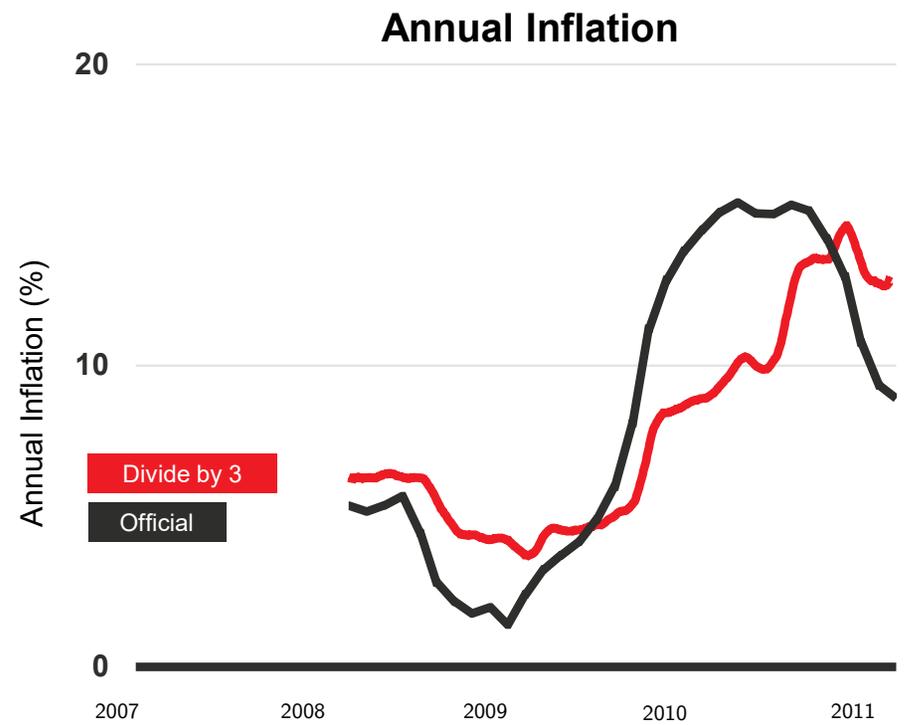
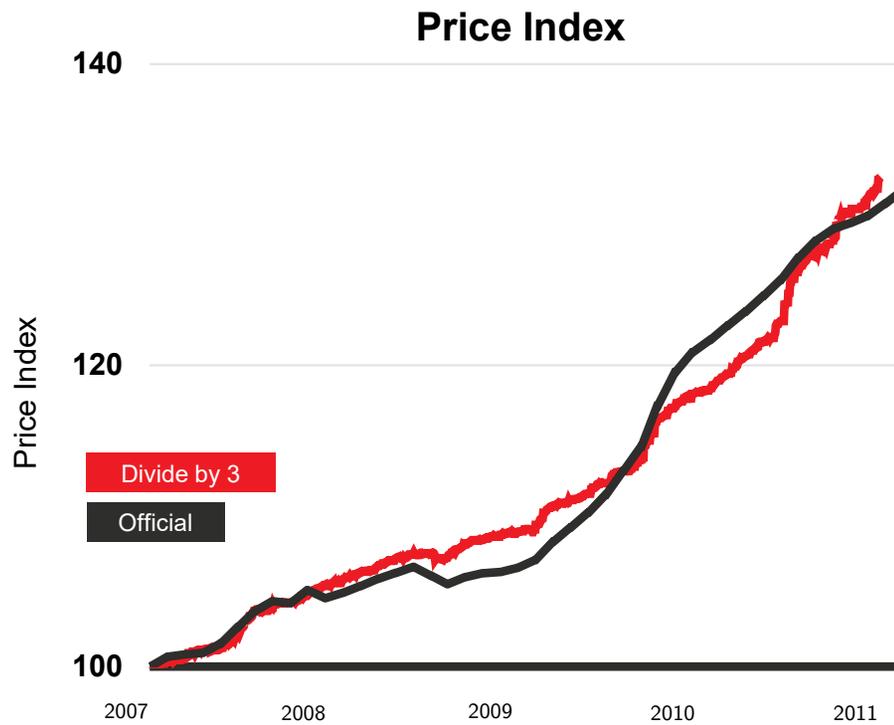
Source: Cavallo (2012) "Online and Official Price Indexes: Measuring Argentina's Inflation", Journal of Monetary Economics

Difference was not explained by data source, method, or type of goods



Source: Cavallo (2013) Online vs Official Price Indexes: Measuring Argentina's Inflation - Journal of Monetary Economics. Vol 60.

WHAT HAPPENS IF WE DIVIDE BY 3?

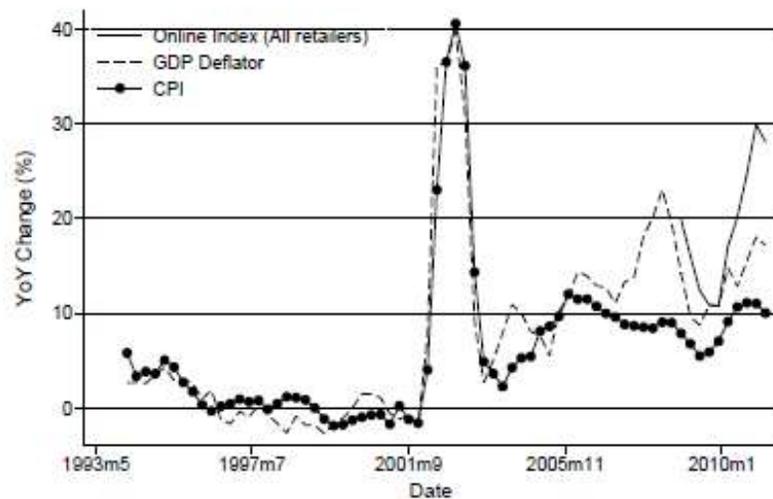


Source: Cavallo (2012) "Online and Official Price Indexes: Measuring Argentina's Inflation", Journal of Monetary Economics

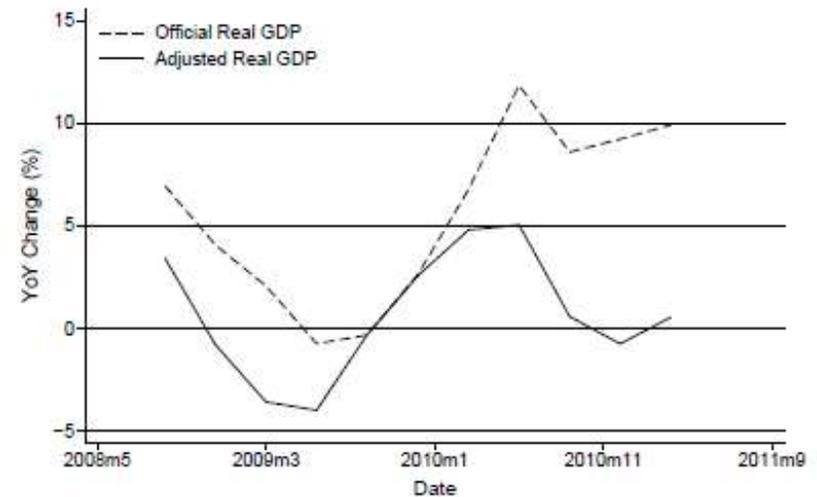
Statistical Uncertainty

- By underestimating inflation they overestimated growth

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



(b) Real GDP - Annual Growth Rate



Todos los días buscamos los precios de un listado de productos en supermercados y calculamos la verdadera inflación de alimentos en Buenos Aires.

Ultima Actualización: 1 Nov 2011 - [Enviar por Email](#) - [Si no podés ver los gráficos instalá el [Flash Player](#)]

Los Precios Hoy

Somos un [grupo de economistas](#) que todos los días registra los precios de un [listado](#) de productos en dos grandes supermercados de Buenos Aires y los publica en este sitio.

En esta sección mostramos un resumen de nuestras principales estadísticas, basadas en la evolución diaria de dos índices propios de inflación:

- Un *índice de Alimentos y Bebidas*, equivalente al 31% del IPC.
- Un *índice de la Canasta Básica Alimentaria*, como el que el INDEC usa (o usaba) para calcular el nivel de indigencia

En la sección [Novedades](#) hacemos un análisis periódico de los datos. En la secciones de [Precios](#), [Aumentos](#) y [Caídas](#) se pueden ver detalles de precios individuales. Finalmente, en [Metodología](#) y [Preguntas](#) explicamos cómo lo hacemos y cómo podés ayudarnos.

Compartir |    



[Novedades](#)

Email:

Administrado por FeedBurner

 Suscribir con RSS

INFLACION ANUAL

23,7 %

Alimentos y Bebidas
www.InflacionVerdadera.com

INFLACION ANUAL

25,4 %

Canasta Basica
www.InflacionVerdadera.com



Fuente: www.InflacionVerdadera.com

Fuente: www.InflacionVerdadera.com

(incluye precios que entran y salen de oferta)

DON'T LIE TO ME, ARGENTINA

“...for the vast majority of Argentines, the indices compiled by INDEC...have been grossly manipulated.”

— La Nación Newspaper, March 10, 2008

DON'T LIE TO ME, ARGENTINA

“...we have decided to drop INDEC's figures entirely...”

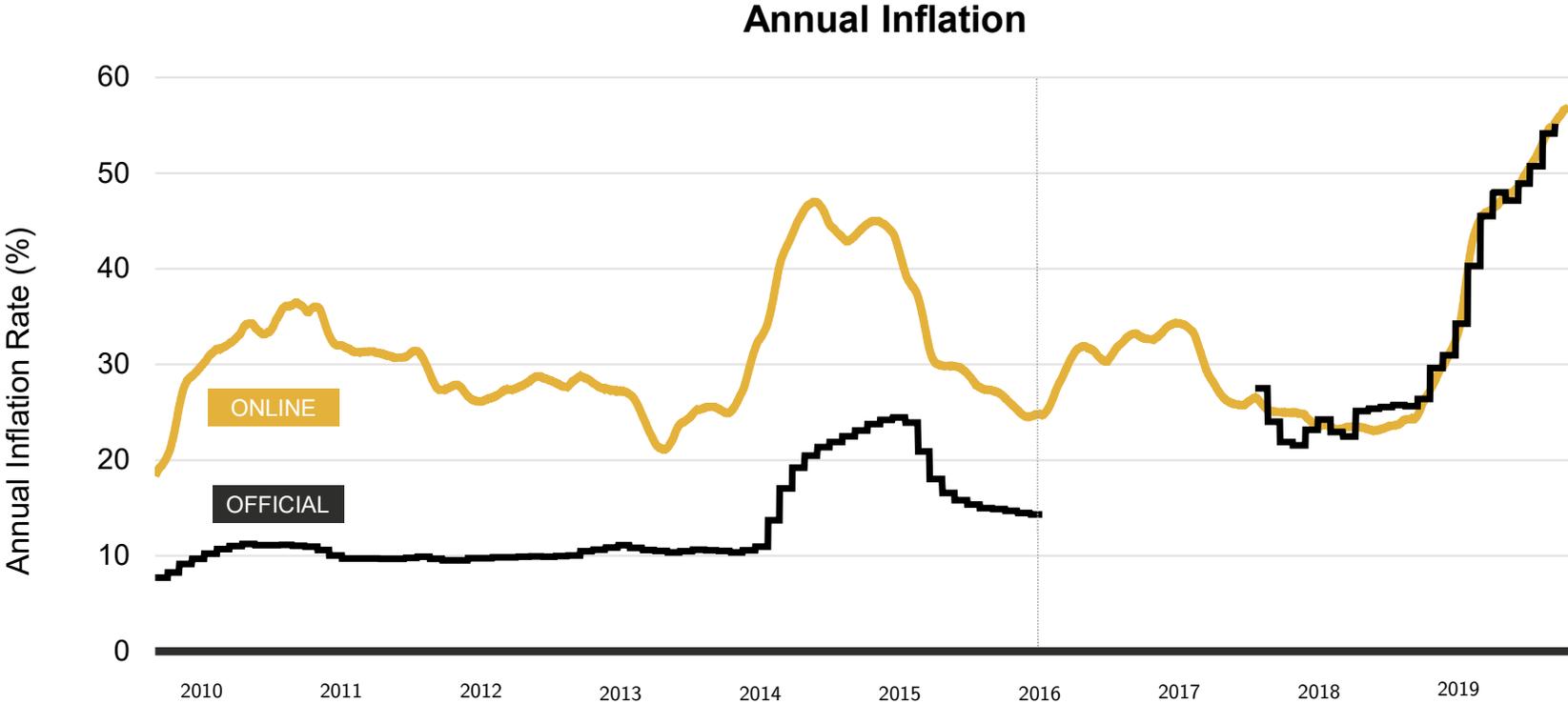
—The Economist, February 25, 2012

DON'T LIE TO ME, ARGENTINA

“... a declaration of censure...the Board called Argentina to...address the inaccuracy of the CPI.”

—IMF, February 2013

EVENTUALLY GOVERNMENT TELLS THE TRUTH



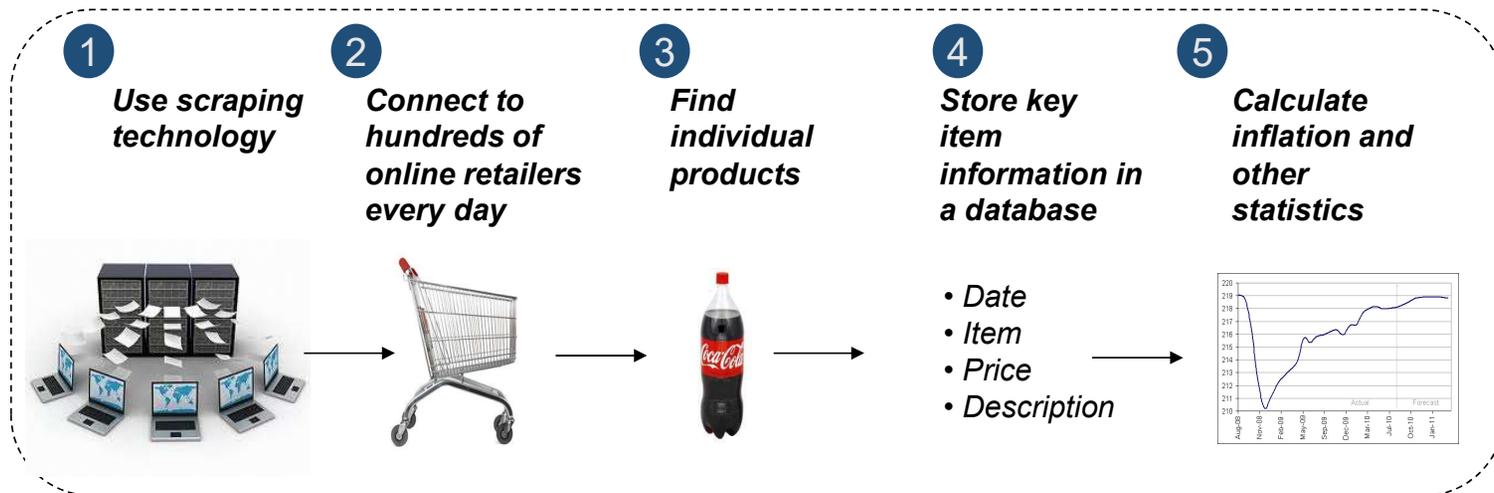
Source: PriceStats, INDEC, The Billion Prices Project

THE BILLION PRICES PROJECT: PROJECT: MEASUREMENT AND RESEARCH



The Billion Prices Project

- Academic initiative to collect and use online price data for economic measurement and research
 - Daily prices since 2008
 - From hundreds of large multi-channel retailers
 - In over 60 countries



www.thebillionpricesproject.com

Micro-Price Data: Advantages and Disadvantages

Table 1: Alternative Micro-Price Data Sources

	Online Data	Scanner Data	CPI Data
Cost per observation	Low	Medium	High
Data Frequency	Daily	Weekly	Monthly
All Products in Retailer (Census)	Yes	No	No
Uncensored Price Spells	Yes	Yes	No
Countries with Research Data*	~60	<10	~20
Comparable Across Countries	Yes	Limited	Limited
Real-Time availability	Yes	No	No
Product Categories Covered	Few	Few	Many
Retailers Covered	Few	Few	Many
Quantities or Expenditure Weights	No	Yes	Yes

Note: Table from Cavallo (2015b). *Approximate numbers. The Billion Prices Project (bpp.mit.edu) datasets contain information from over 60 countries with varying degrees of sector coverage. Nielsen US scanner datasets are available at the Kilts Center for Marketing of the University of Chicago. Klenow and Malin (2010) provide stickiness results with CPI data sourced from 27 papers in 23 countries. See Cavallo (2013) for more details.

Source: Cavallo & Rigobon (2016) "The Billion Prices Project", *Journal of Economic Perspectives*, Spring 2016, Vol 30(2):151-78.

The Billion Prices Project - Research

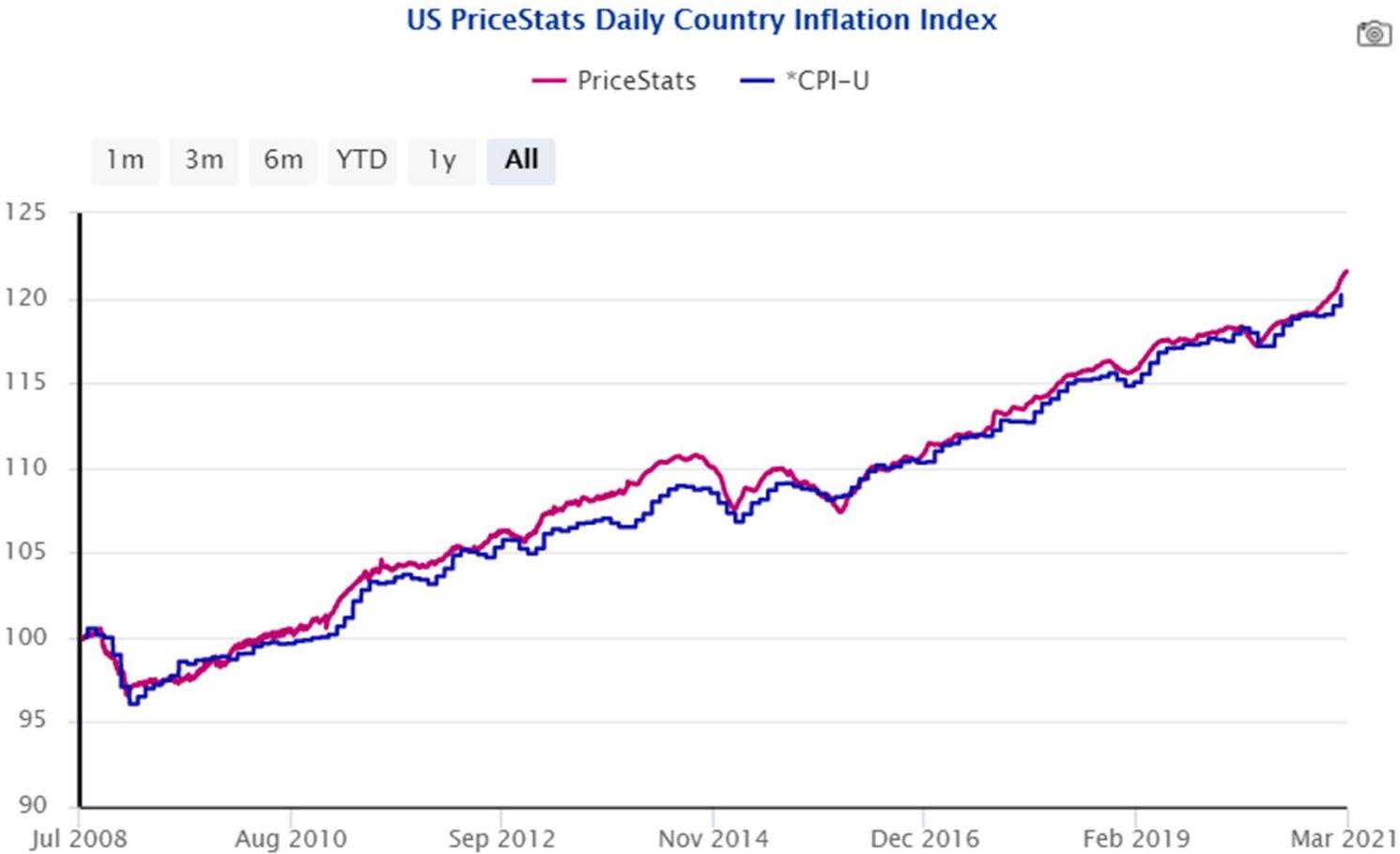
Research Topic	Papers
Macroeconomics	
Inflation Measurement	Cavallo (2020a)
Online Price Indices	Cavallo (2013), Cavallo & Rigobon (2016)
Crowdsourcing with phones in Venezuela	Cavallo (2020b)
Online Pricing and Price Discrimination	Cavallo (2017), Cavallo (2018a)
Price Stickiness	Cavallo (2018b)
Inflation Expectations	Cavallo, Cruces, Perez-Truglia (2016,2017)
International Economics	
LOP and market segmentation	Cavallo, Neiman, Rigobon (2014,2015)
Purchasing Power Parities	Cavallo, Diewert, Feenstra, Inklaar & Timmer (2018)
International Wage Comparisons	Cavallo, Cravino, & Drenik (2019)
Tariff passthrough	Cavallo, Neiman, Gopinath, & Tang (2020)

Daily Inflation Measurement

- 2008 → Daily index for Argentina (www.inflacionverdadera.com)
- 2010 → Daily index for the US on the BPP website
- 2011 - now → PriceStats collects the micro data and publishes daily inflation in 23 countries in real-time (3-day lag).



Daily Inflation Measurement

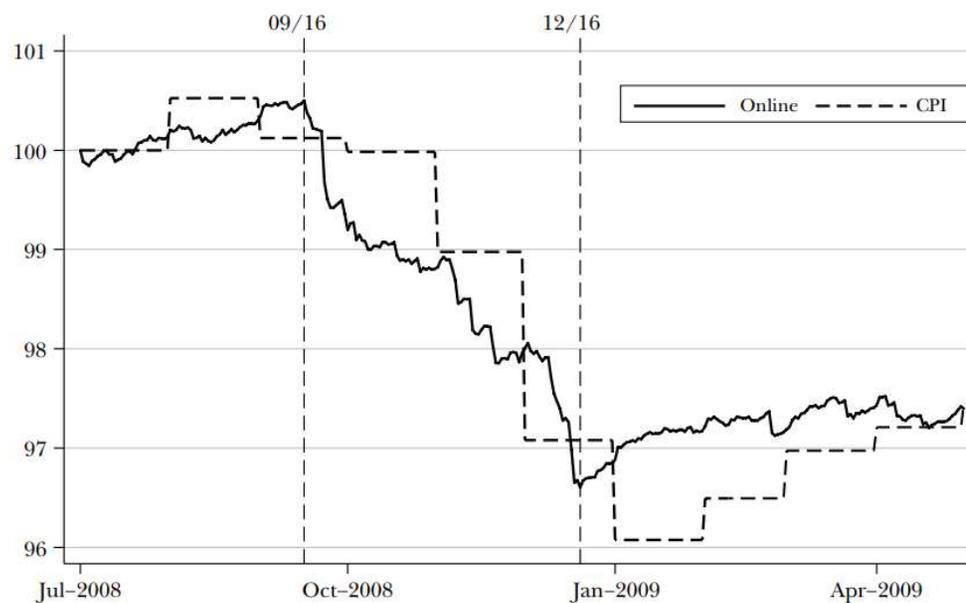


Source: PriceStats - State Street Global Markets Research, and BLS

Daily Inflation Measurement

Figure 3

US Consumer Price Index around the Bankruptcy of Lehman Brothers

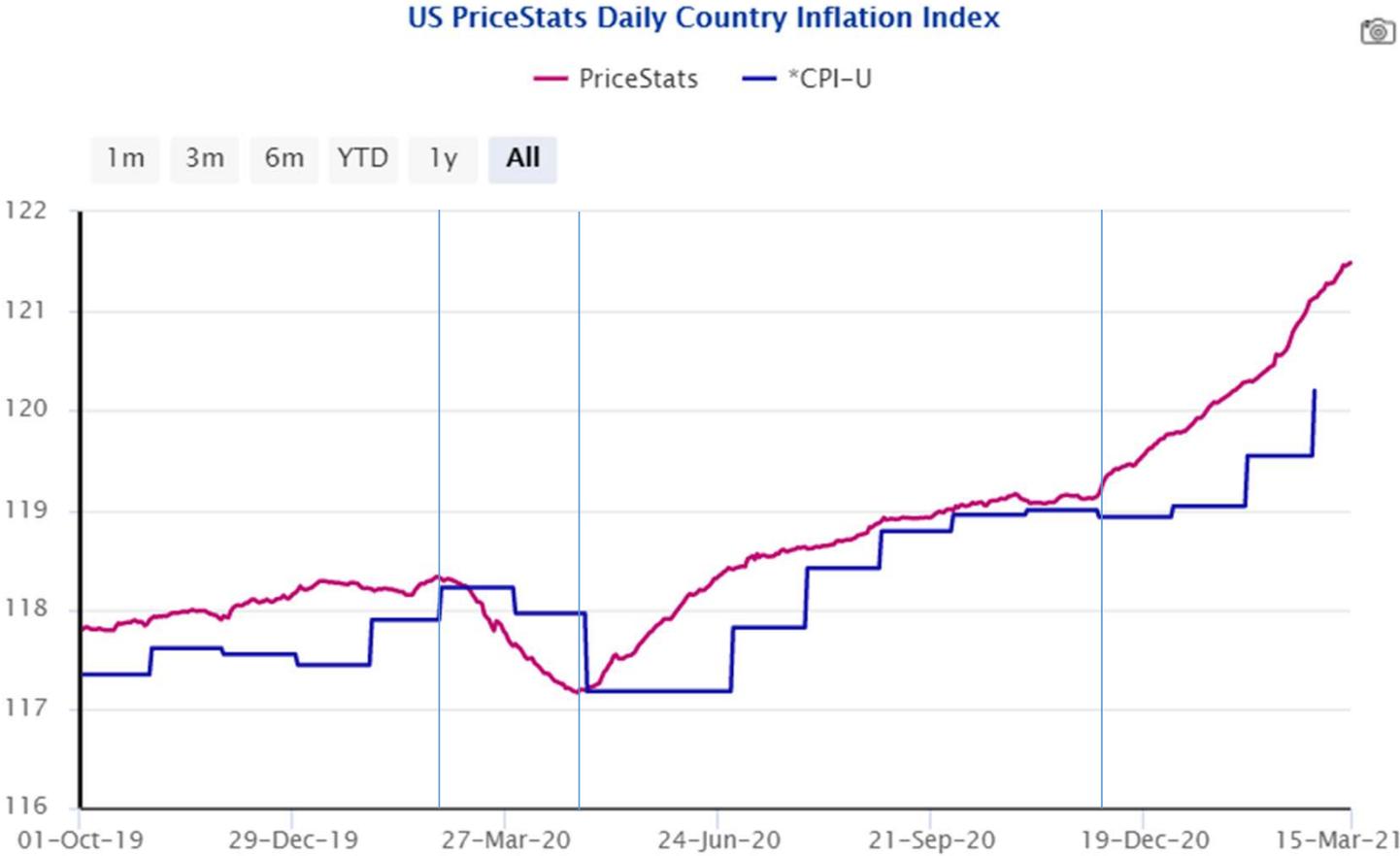


Source: Authors using online price index computed by PriceStats and the Consumer Price Index from the US Bureau of Labor Statistics.

Note: The figure highlights the events around the bankruptcy of Lehman Brothers, the fourth-largest investment bank in the United States, during September 2008.

Source: Cavallo & Rigobon (2016) "The Billion Prices Project", Journal of Economic Perspectives, Spring 2016, Vol 30(2):151-78.

Daily Inflation Measurement

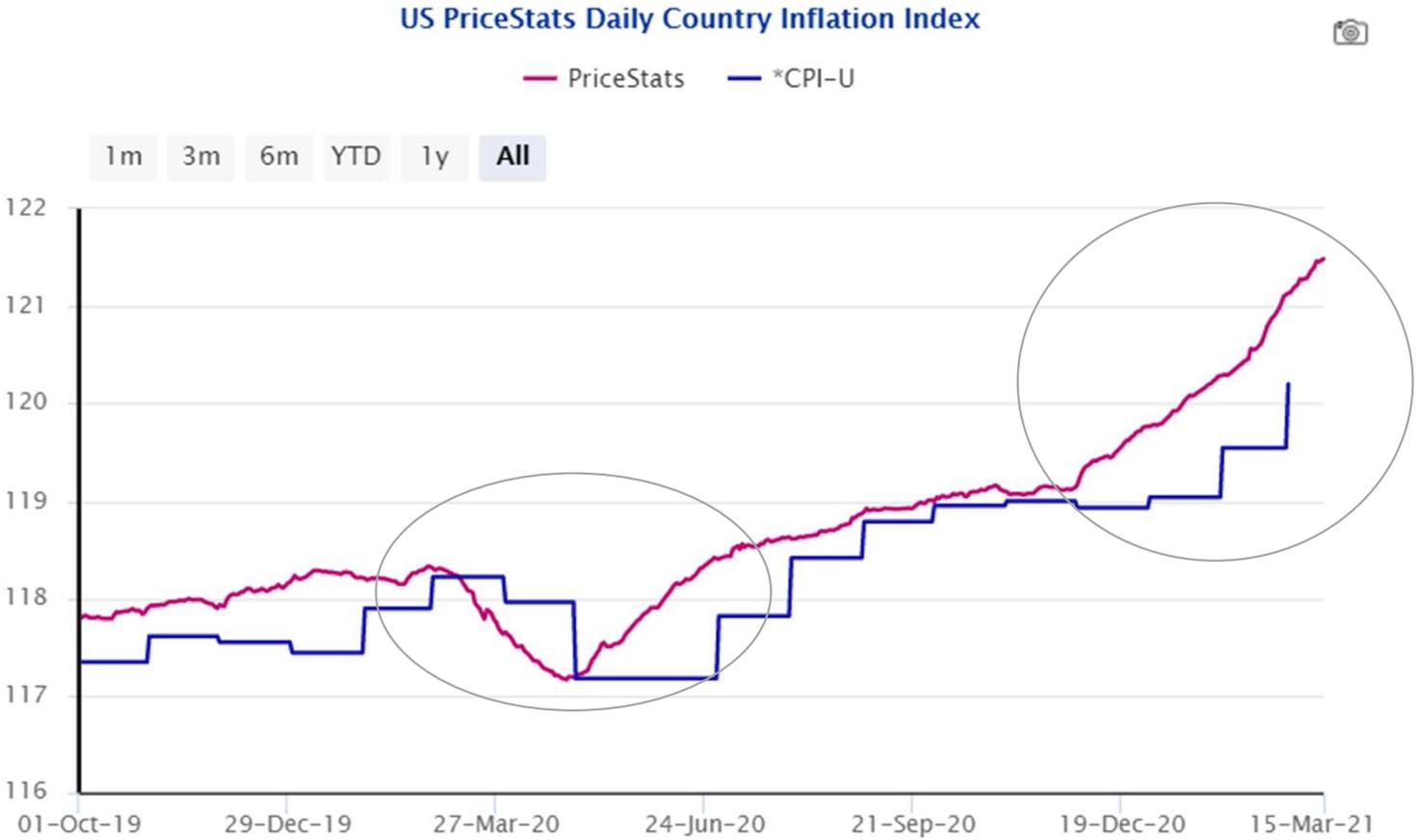


Source: PriceStats - State Street Global Markets Research, and BLS

Applications in Forecasting

- Bertolotto (2019), Aparicio & Bertolotto (2020)
- Advantages for forecasting:
 - Frequency (daily) and speed of publication (3-day lag)
 - Provides information during crises and turning points , when traditional forecasting models based on lagged CPIs and low-frequency data tend to fail the most
- Real-time measurement helps to nowcast the present (1-2 months ahead)
- But for longer-term forecasting → we need to better understand shocks and inflation dynamics

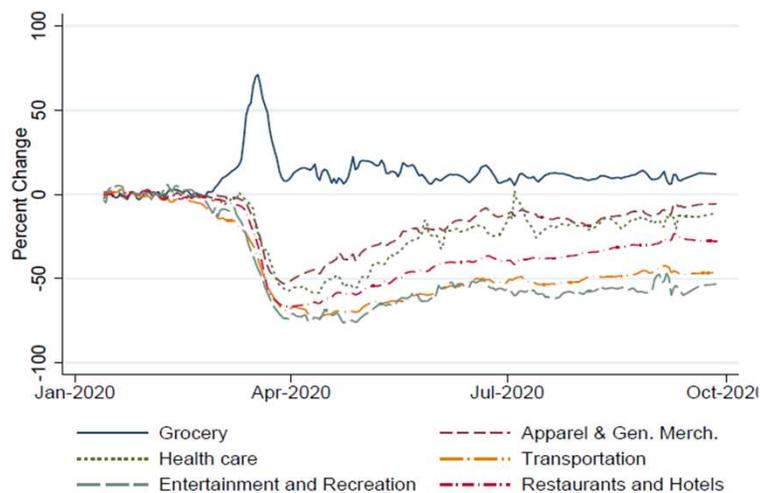
Covid Inflation Dynamics



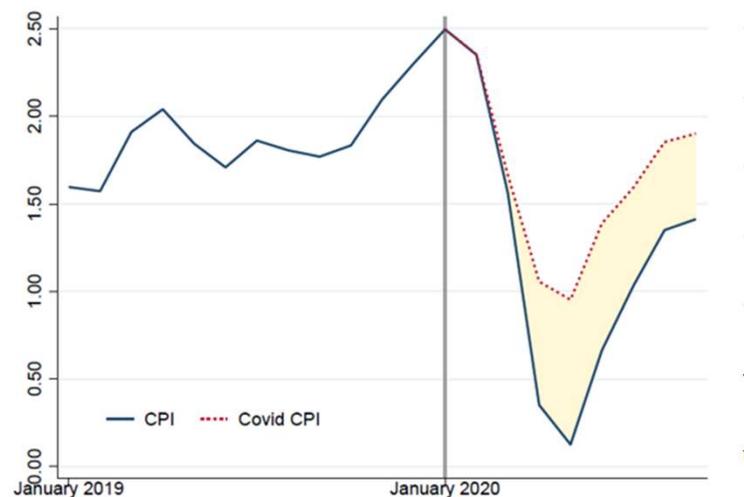
Source: PriceStats - State Street Global Markets Research, and BLS

Research related to Covid Inflation Dynamics

- Covid inflation was **higher** than measured by the CPI (Cavallo (2020) Inflation with COVID Consumption Baskets . NBER Working Paper No. 27352)



(a) Consumer Spending (Opportunity Insights)



(b) Annual Inflation Rate (%)

Source: Cavallo, Alberto. "Inflation with Covid Consumption Baskets." NBER Working Paper Series, No. 27352, July 2020. The consumer spending time series are computed by Chetty et al (2020) using credit and debit card transactions. The data are publicly available at the Opportunity Insights website (opportunityinsights.org).

Updated results, all data and codes are available at) <https://projects.iq.harvard.edu/covid-cpi>

Research related to Covid Inflation Dynamics

- Covid inflation was **higher** than measured by the CPI (Cavallo (2020) Inflation with COVID Consumption Baskets . NBER Working Paper No. 27352)
- Online prices tend to **react faster** to aggregate shocks (Cavallo (2018) More Amazon Effects: Online Competition and Pricing Behaviors , Jackson Hole Symposium)
 - More frequent price changes + uniform pricing → more cost-shock passthrough

Monthly Frequency of Price Changes, 2008 to 2017

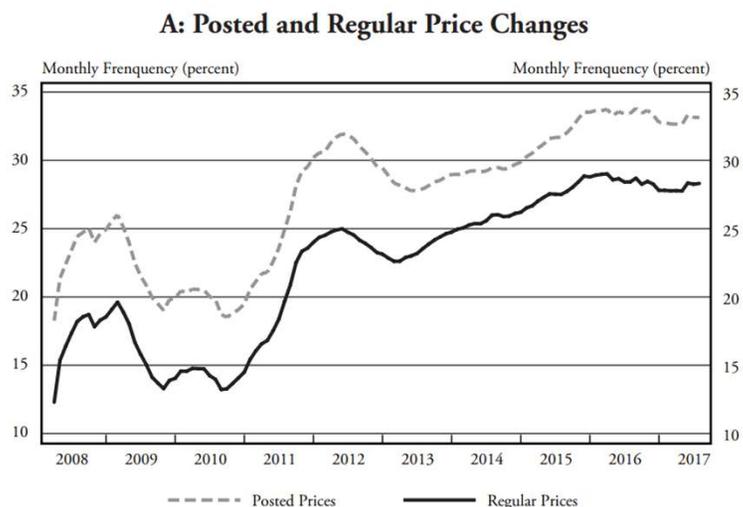


Table 6: Short-Run Pass-through into Walmart's Prices (2016-2018)

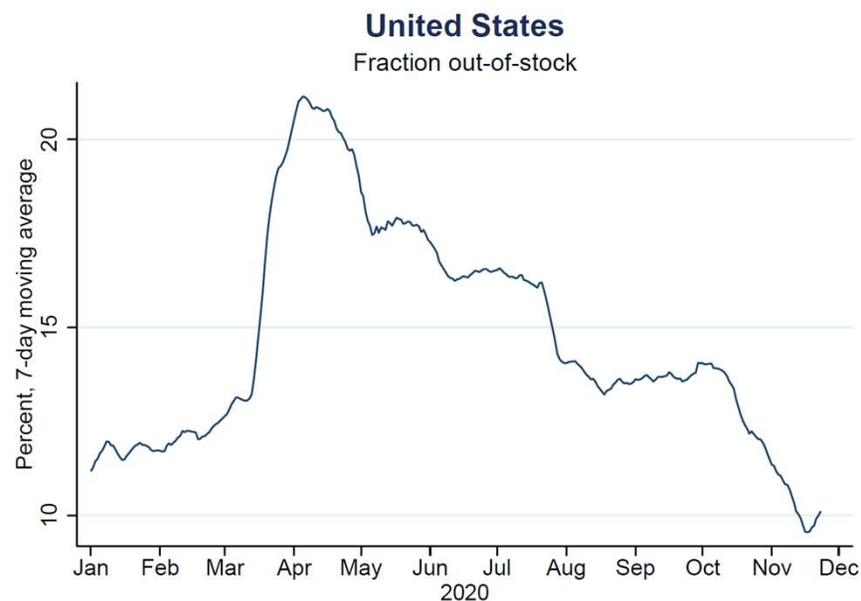
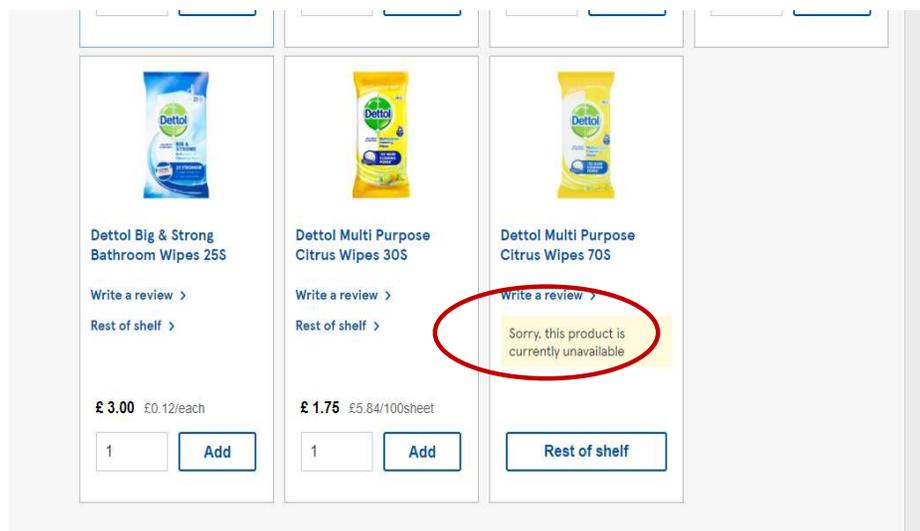
	Full Sample	Found on Amazon	
		No	Yes
Gas Prices (1 quarter)	0.22 (0.02)	0.19 (0.02)	0.28 (0.03)
Observations	191,690	122,800	68,890
R-squared	0.17	0.17	0.16
Exchange Rate (2 quarters)	0.32 (0.03)	0.26 (0.04)	0.44 (0.05)
Observations	191,690	122,800	68,890
R-squared	0.17	0.18	0.16

Research related to Covid Inflation Dynamics

- Covid inflation was **higher** than measured by the CPI (Cavallo (2020) Inflation with COVID Consumption Baskets . NBER Working Paper No. 27352)
- Online prices tend to **react faster** to aggregate shocks (Cavallo (2018) More Amazon Effects: Online Competition and Pricing Behaviors , Jackson Hole Symposium)
 - More frequent price changes + uniform pricing → more cost-shock passthrough
- Retail cost-passthrough can be slow when shock is **perceived to be temporary**, and firms have **other ways to adapt** (Cavallo et al (2020) Tariffs Passthrough at the Border and at the Store: Evidence from US Trade Policy AER Insights)
- Covid **supply disruptions** are putting upward pressure on prices (Cavallo & Kryvtsov (2021) Stockouts and Prices during the Covid-19 Pandemic , Working Paper)

Supply Disruptions and Stockouts

- We measure product stock-outs in 7 countries for a subset of retailers showing out of stock information

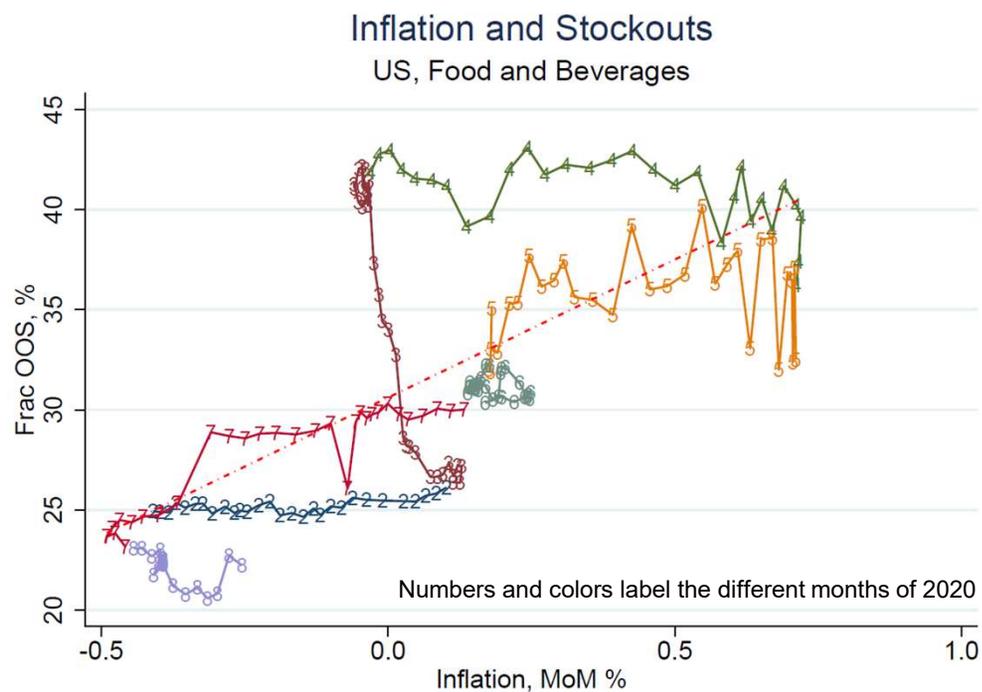


Note: Aggregate-level stockouts using CPI category weights

Source: Cavallo & Kryvtsov (2021) "Stockouts and Prices during the Covid-19 Pandemic", Working Paper (preliminary results). The views expressed here are ours, and they do not necessarily reflect the views of the Bank of Canada.

Stockouts and Prices

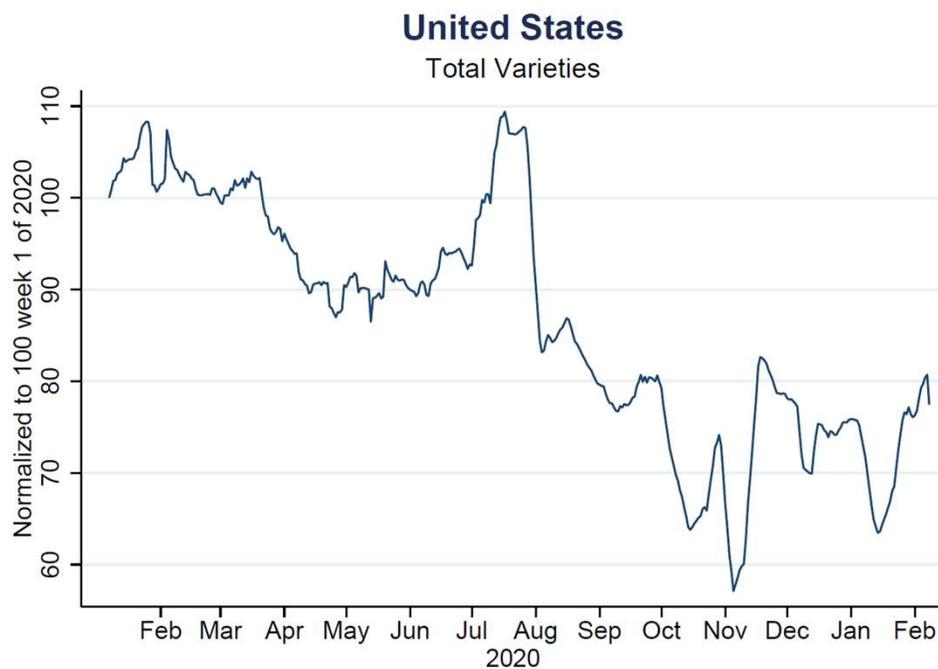
- Stockouts were positively correlated with US sectoral inflation in 2020



Source: Cavallo & Kryvtsov (2021) "Stockouts and Prices during the Covid-19 Pandemic", Working Paper (preliminary results). The views expressed here are ours, and they do not necessarily reflect the views of the Bank of Canada.

Product availability remains low

- Stockouts fell but the total number of varieties available for sale at these retailers is still ~20% below pre-pandemic levels



→ supply disruptions are still important in many retail categories and may continue to put upward pressure on prices for months

Source: Cavallo & Kryvtsov (2021) "Stockouts and Prices during the Covid-19 Pandemic", Working Paper (preliminary results). The views expressed here are ours, and they do not necessarily reflect the views of the Bank of Canada.

Conclusions

- Online data provide a unique measurement opportunity
 - Speed, frequency, details & customization, alternative to official sources
- Real-time price indices are useful for nowcasting inflation, particularly during times of crisis and shocks
- More importantly, better micro data can greatly improve our understanding of shock pass-through and longer-term inflation dynamics

Additional Slides

Are Online and Offline Prices Similar?

- Cavallo (2017) Are Online and Offline Prices Similar: Evidence from Large Multi-Channel retailers , American Economic Review
- Large-scale comparison of online and offline prices collected simultaneously in ~50 retailers in 10 countries.
- Crowdsourced workers scan random barcodes, enter prices, send emails with data files.



App available for download at the Google Play Store: <https://play.google.com/store/apps/details?id=com.mit.bpp>

Figure 1: Screenshots from BPP App for Android Phones

- We then scraped the online price for the same good-retailer (within 7 days).

Prices are identical ~70% of the time

Table 3: Country - Level Differences

Country	(1) Ret.	(2) Obs	(3) Identical (%)	(4) High On (%)	(5) Low On (%)	(6) Markup (%)	(7) Difference (%)
Argentina	5	3699	60	27	13	3	1
Australia	4	3797	74	20	5	5	1
Brazil	5	1915	42	18	40	-7	-4
Canada	5	4031	91	3	5	-5	0
China	2	513	87	7	6	3	0
Germany	5	1604	74	4	23	-8	-2
Japan	4	2186	48	7	45	-13	-7
South Africa	5	3212	85	6	9	-3	-1
UK	4	2094	91	2	7	-8	-1
USA	17	15332	69	8	22	-5	-1
ALL	56	38383	72	11	18	-4	-1

Note: Results updated 5 Apr 2016. Column 3 shows the percentage of observations that have identical online and offline prices. Column 4 has the percent of observation where prices are higher online and column 5 the percentage of price that are lower online. Column 6, is the online markup, defined as the average price difference excluding cases that are identical. Column 7 is the average price difference including identical prices.

Table 4: Sector - Price Level Differences

Sector	(1) Ret.	(2) Obs	(3) Identical (%)	(4) High On (%)	(5) Low On (%)	(6) Markup (%)	(7) Difference (%)
Food	10	5953	52	32	15	3	1
Clothing	7	2534	92	5	3	3	0
Household	9	7875	79	5	16	-8	-2
Drugstore	4	3053	38	11	52	-5	-3
Electronics	5	3712	83	4	13	-9	-1
Office	2	1089	25	37	38	1	1
Multiple/Mix	18	14149	80	5	15	-9	-2

Note: Results updated 5 Apr 2016. Markup excludes identical prices. Difference includes identical prices.

Practice some web scraping

- It is not as hard as you think:
 - All webpages use an HTML code with tags that provide a stable structure to identify the data → you can teach a software to recognize them
 - Many alternative tools: R, python, specialized software, scraping services
 - Similar Steps:
 - 1) Create a template for parsing the HTML code
 - 2) Create a list of URLs with the relevant data
 - 3) Run scraper, analyze data, debug
- Not just for retail prices:
 - Wage postings (Cavallo, Cravino & Drenik 2019), stockouts (Cavallo & Kryvtsov 2020), real estate listings, customs data, wholesale prices, or any structured data that can be found online

US Annual Inflation

US PriceStats Daily Country Inflation Index (% yoy)



PriceStats *CPI-U

1m 3m 6m YTD 1y All



References

- Aparicio, D., Bertolotto, M. 2020: Forecasting Inflation with Online Prices. *International Journal of Forecasting*, Vol 36, Issue 2.
- Bertolotto, M., 2019: Online Price Indices: Methodological Improvements and Applications to Inflation Forecasting and RER Estimation. PhD thesis, UdeSA, Argentina.
- Cavallo, A., 2013. Online and official price indexes: Measuring Argentina's inflation. *Journal of Monetary Economics* 152-165.
- Cavallo, A., 2017. Are Online and Offline Prices Similar? Evidence from Large Multi-Channel Retailers. *American Economic Review* 107.
- Cavallo, A., 2018a. More Amazon Effects: Online Competition and Pricing Behaviors. Jackson Hole Economic Symposium Conference Proceedings - Federal Reserve Bank of Kansas City.
- Cavallo, A., 2018b. Scraped Data and Sticky Prices. *Review of Economics and Statistics*, vol 100-1.
- Cavallo, A., 2020a. Inflation with Covid Consumption Baskets. NBER Working Paper 27352.
- Cavallo, A., 2020b. Measuring Venezuela's Inflation Using Crowdsourcing and Mobile Phones. Working Paper.
- Cavallo, A., Cravino, J., Drenik, A., 2019. Using Online Data for International Wage Comparisons. HBS Working Paper 20-042.
- Cavallo, A., Cruces, G., Perez-Truglia, R., 2017. Inflation Expectations, Learning, and Supermarket Prices: Evidence from Survey Experiments. *American Economic Journal: Macroeconomics* 9, 1-35.
- Cavallo, A., Cruces, G., Perez-Truglia, R., 2016. Learning from Potentially Biased Statistics: Household Inflation Perceptions and Expectations in Argentina. *Brookings Papers on Economic Activity* 59-108.
- Cavallo, A., Diewert, W.E., Feenstra, R.C., Inklaar, R., Timmer, M.P., 2018. Using Online Prices for Measuring Real Consumption across Countries. *AEA Papers and Proceedings* 108, 483-487.
- Cavallo, A., Gopinath, G., Neiman, B., Tang, J., 2020. Tariff Passthrough at the Border and at the Store: Evidence from US Trade Policy. *American Economic Review: Insights*.
- Cavallo, A., Kryvtsov, O., 2021. Stockouts and Prices During the Covid-19 Pandemic: Evidence from Online Micro Data. Working Paper.
- Cavallo, A., Neiman, B., Rigobon, R., 2015. The Price Impact of Joining a Currency Union: Evidence from Latvia. *IMF Econ Rev* 63, 281-297.
- Cavallo, A., Neiman, B., Rigobon, R., 2014. Currency Unions, Product Introductions, and the Real Exchange Rate. *Quarterly Journal of Economics* 129.
- Cavallo, A., Rigobon, R., 2016. The Billion Prices Project: Using Online Data for Measurement and Research. *Journal of Economic Perspectives* 30, 151-78.
- Eichengreen, B., 2015. *An Economics to Fit the Facts*, Project Syndicate.
- Einav, L., Levin, J., 2014. Economics in the age of big data. *Science* 346.
- Griliches, Z., 1985. Data and econometricians: the Uneasy Alliance. *The American Economic Review* 196-200.
- Summers, L.H., 2020. Data collection is the ultimate public good. *Washington Post*. public good. *Washington Post*.