

# Online and Official Price Indexes: Measuring Argentina's Inflation

Alberto Cavallo<sup>1,\*</sup>

*MIT Sloan*  
*February 27, 2012*

---

## Abstract

Prices collected from online retailers can be used to construct high-frequency price indexes that complement official statistics. This paper studies their ability to match official inflation estimates in five Latin American countries, with a focus on Argentina, where official statistics have been criticized in recent years. The data were collected between October 2007 and March 2011 from the largest online supermarkets in each of these countries. Online price indexes approximate both the level and main dynamics of official inflation in Brazil, Chile, Colombia, and Venezuela. By contrast, Argentina's online annual inflation rate is consistently two to three times higher than in official estimates.

---

---

\*Address: 77 Mass Ave E62-512, Cambridge MA 02139 - [acavallo@mit.edu](mailto:acavallo@mit.edu) - <http://acavallo.mit.edu>

<sup>1</sup>An early version of this paper, titled "Argentina: Inflation and Lies", was part of my dissertation at Harvard University. I thank my advisors, Philippe Aghion, Robert Barro, and Roberto Rigobon, for their support and encouragement. I am also grateful to Barry Bosworth, Erwin Diewert, Ernie Berndt, Brian Kovak, and seminar participants at Harvard, MIT, NBER, and LACEA; to the editor and two referees, for their helpful comments and suggestions.

JEL Code: E31.

Keywords: Argentina, inflation, scraped online data, online price index

## 1. Introduction

The availability of online prices represents a unique opportunity for the construction of price indexes and the measurement of inflation around the world. An unprecedented amount of micro-level price data can now be collected using special software that finds and aggregates detailed product information available in online retailers across the web. This type of data collection can take place at much higher frequencies and a tiny fraction of the cost of traditional methods. Among its many potential uses, one of the simplest applications of online price indexes is to obtain alternative inflation series in countries where official estimates have lost credibility.

This paper uses online prices to evaluate the widespread claim that the Argentine government has been manipulating official inflation indexes since 2007.<sup>2</sup> Online indexes are first shown to be able to approximate both the level and dynamic behavior in inflation trends in four Latin American countries: Brazil, Chile, Colombia, and Venezuela. In Argentina, by contrast, there are huge differences in the level of online and official inflation rates that have persisted for over 3.5 years. A series of robustness tests show that there is no simple data or methodological explanation that can account for this large discrepancy between online and official data.<sup>3</sup>

The online data were collected on a daily basis, between October 2007 to March 2011, by the Billion Prices Project (BPP) at MIT. Every day, a software scanned the websites of some of the largest supermarkets in each of these countries, collecting product-level information and storing it in a database. Over time, a panel database was constructed with detailed information on each product, including prices, product IDs, and a category indicator. Data from six supermarkets are used in this paper, two for Argentina and one for each of the other countries, with an average of 20,752 individual products per retailer.

---

<sup>2</sup>See [1] and [7].

<sup>3</sup>Additional robustness tests are provided in the online Appendix, available in the electronic version of this Journal and the author's academic Web site.

An online price index is built in each country using a combination of online prices, standard CPI methodologies, and official category weights. Each index is then compared to an equivalent official supermarket index, constructed using only the CPI components of Food, Beverages, and Household Products –the same categories available in the online data–.

In Brazil, Chile, Colombia, and Venezuela, the online indexes are able to approximate both the average level and main dynamics of official inflation. The matching is best in Chile, with an average annual inflation of 3.00% online and 3.19% offline, and a correlation of 0.97 in the annual inflation series. The match is also good in Colombia, both for the level and dynamics of annual inflation, which dropped in both online and official series from about 8% in late 2008 to about 3% in late 2009. In Brazil and Venezuela, the online index is able to match the main inflation trend of the official index, but the annual inflation series are less synchronized over time. Not surprisingly, the degree of matching appears to be driven by the representativeness of the data. The supermarket in Chile has a market share of 27%, and Santiago, the city where the prices are collected, represents 55% of the national CPI. By contrast, the supermarket in Brazil has just a 15% market share, and online prices are collected in Rio, which represents only 17.3% of the national CPI.

The results for Argentina are remarkably different. There is a large discrepancy between online and official price indexes that is persistent over time. After three and a half years, the online index grew over 100%, while the official index increased by only 35%. Online prices had an annual inflation rate that was consistently two to three times larger than in official statistics, with an average rate of 20.14% for the period, compared to just 8.38% in official data. Surprisingly, although the *level* of inflation is higher, the dynamic behavior of inflation matches the official data quite well, both at the annual and monthly frequencies. In particular, both the online and official series show a decrease in annual inflation rates during 2009, when the economy was going through a recession.

The results in Argentina are puzzling because there are reasons to expect the online data to be more representative than other countries. The Argentine supermarket has a large

market share and targets consumers with a broad range of income levels. Furthermore, the online prices come from Buenos Aires, which is also the only location where official data collection takes place.

Alternative methodologies and subsets of the data yield similar results. First, an online index is built using data from a second supermarket in Argentina with completely different characteristics: a small, online-only retailer that targets high-income people in Buenos Aires. Second, an index is constructed exclusively with goods that have price controls. Finally, a simple “Subsistence Food” Index, with only 45 goods, is compared to the “Canasta Básica de Alimentos” (CBA) that INDEC uses to calculate the level of extreme poverty in the country. In all cases, the main results are the same, with large unexplained differences with the equivalent official series. In fact, the only way to approximate the official index in Argentina is to use one-third of the inflation rate observed online.

These results confirm the widespread suspicion that the government has been manipulating the consumer price index since January 2007, when it intervened the National Statistical Institute (INDEC). The implications for other statistics are significant. For example, using an online-inflation adjusted cost for the subsistence-level CBA basket, the share of the population in extreme poverty during the first quarter of 2011 would rise to 6.69%, instead of only 2.5% as officially reported. Similarly, poverty shares would be 25.9% instead of the 9.9% in official estimates. The implications for real GDP can be equally important. If the GDP deflator had behaved like the online index since 2007, the real GDP growth rate would have been just 0.5% in March 2011, instead of 10% officially reported.

The paper is organized as follows. Section 2 describes the data. Section 3 provides the methodology for online price indexes. Section 4 presents the results for Brazil, Chile, Colombia, and Venezuela. Section 5.2 discusses the case of Argentina. Section 6 concludes with a discussion of why the policy implications, as well as some opportunities to use online price indexes in other countries.

## 2. The Data

The data were collected by the Billion Prices Project (BPP) at MIT using a technique called “web scraping” to record the price for all goods sold online, between October 2007 and March 2011, in the largest supermarket in Argentina, Brazil, Chile, Colombia, and Venezuela.

The technology to “scrape” prices is conceptually simple. Most webpages are built using a structured coding language called HyperText Markup Language (HTML). This code has simple “tags”, such as `<center>` and `<bold>`, that determine the style and placement of text in a given page. Unless the retailer decides to change the “look and feel” of its Webster, these tags tend to remain constant over time. The information *within* these tags, such as a product’s price, can change at any time. The scraping software can be taught to use the HTML tags to locate relevant information about a product, such as its price, and store it in a database. Repeating the process every day produces a panel database with a one record per product per day. Retailers usually put a product ID in the HTML code, which is constant over time and can be used to build time series of individual prices. Finally, the web address or “URL” of the page where each product is located can be used to classify products into standardized categories.<sup>4</sup>

Table 1 describes the six databases included in this paper: two for Argentina, and one for each of the other countries. Argentina’s Retailer #1 is the largest supermarket chain in the country. It is used for the main online price index because it has detailed category indicators and large market share. Retailer #2, used for robustness in section 5.3.2, is a smaller supermarket that sells exclusively through the web to a higher-income population. The supermarkets in Brazil, Chile, and Colombia are the market leaders, with websites that target people in the cities of Rio de Janeiro, Santiago, and Bogota. The supermarket in Venezuela is a smaller store that targets mostly the high-income population of Caracas.

In all cases, the online data contains a combination food, beverages, and household prod-

---

<sup>4</sup>For example, all products shown at the URL “www.supermarketname.com/fatfree-milk” would be classified as “Milk”.

ucts. Categories range from “Eggs” to “Appliances”, with about a third of them corresponding to household products, including cleaning materials, health and beauty products, furniture, appliances, and books. Together these categories account for between 28.44% (Colombia) and 48.51% (Argentina) of CPI weights in these countries, as shown in Table 1.<sup>5</sup>

### 3. Online Price Indexes

Online price indexes use a combination of online data and official category weights. The methodology follows the way CPI statistics are constructed in these countries as closely as possible, so that the main difference is the source of the price information. There are, however, some differences in the treatment of the data. First, online indexes are computed with daily data. A daily index is useful to observe short-term patterns in the data that help validate the online information, as discussed in Section 5.3.3. Second, the online indexes are constructed using all products on display at each retailer. Within these particular good categories, the sample tends to be much larger than in official statistics, and also the basket of goods changes dynamically over time, as products appear or disappear from the online stores. Results are robust to the use a small fixed basket of goods, as shown in Section 5.3.3. Third, there are no forced product substitutions or adjustment for quality changes. All goods are treated independently, so products that are discontinued on a given date will stop affect the index from that day forward. Similarly, new goods will only impact the index on their second day in the sample, when their first price change can be calculated. This is not a problem in this paper, because substitutions and quality adjustments are not common in official statistics for the categories of goods analyzed in this paper. Finally, with daily data, it is common to find short gaps in an individual price series, caused either by failures in the scraping method or products that go temporarily out of stock. In this paper, the gaps are filled by carrying forward the last available price for each product, but all results are robust

---

<sup>5</sup>Detailed official category weights, for the products available in each supermarket, are shown in Appendix Table B.1.

to the use of cell-relative adjustments.<sup>6</sup>

To build the index, price changes are calculated at the product level, then averaged inside categories using unweighted geometric means, and finally aggregated across categories with a weighted arithmetic mean. In particular, the first step is to obtain the unweighted geometric average of price changes in category  $j$  for each day  $t$ :

$$R_{t,t-1}^j = \prod_i \left( \frac{p_t^i}{p_{t-1}^i} \right)^{\frac{1}{n_{j,t}}} \quad (1)$$

where  $p_t^i$  is the price of good  $i$  at time  $t$ ,  $n_{j,t}$  is the number of products in category  $j$  that are present in the sample that day.

The second step is to compute the category-level index at  $t$ :

$$I_t^j = R_{1,0}^j \cdot R_{2,1}^j \cdots R_{t,t-1}^j \quad (2)$$

Finally, the Supermarket Index at time  $t$  is the weighted arithmetic average of all category indexes:

$$S_t = \sum_j \frac{w^j}{W} I_t^j \quad (3)$$

where  $w^j$  is the official CPI weight for that category and  $W$  is the sum of all the weights included in the sample.

The classification of products and weighing of categories is one of the most complex parts of this process. In the original data, each product is linked to a web address (URL) that corresponds to the webpage where the product is located. These URLs group similar items together, into various levels of aggregation chosen by each supermarket. The number of URLs ranges from about 300 to 1000 in these retailers. The advantage of having URLs is that thousands of items can be easily classified into a set of standardized official categories,

---

<sup>6</sup>Cell-relative adjustment are used by the Bureau of Labor Statistics to fill price gaps in the US CPI. See Appendix B.2 for details.

so that their corresponding official weights  $w^j$  could be used to obtain the aggregate index.<sup>7</sup>

After computing the price index, daily estimates for the “monthly” and “annual” inflation rates are obtained. The monthly inflation rate measures the percentage change in the average index of the last 30 days with respect to the average of the previous 30 days. For example, on November 30th 2010 it is the percentage change between the average of the daily index from November 1st to November 30th 2010, and the average of the daily index from October 2nd to October 31st 2010. Similarly, the annual inflation rate is the percentage change in the average index of the last 30 days with respect to the average of the same period a year ago.<sup>8</sup>

Compared to CPI statistics, these online indexes have an advantage in terms of frequency and the number of items sampled within each category. For example, in Argentina alone, there are prices for 781 different product varieties in the “Milk” category alone (this includes different brands and package sizes). A main disadvantage is that these online prices come from a single retailer in only one city. As discussed below, this can limit the ability of online indexes to match the short-term inflation dynamics. However, this problem is minimized in this case because these Latin American supermarkets have huge market shares, shown in Table 1, and are located in cities that concentrate a large percentage of each country’s population and CPI weights.<sup>9</sup>

#### 4. Four Latin American Countries

This section studies the performance of online price indexes in Brazil, Chile, Colombia, and Venezuela. These countries have similar degrees of internet development as Argentina,

---

<sup>7</sup>Table B.1 of the Appendix has a detailed list of categories and weights used. A complete list of official weights in each country is available in [6], [9], [10], [4], and [13].

<sup>8</sup>The annual inflation rate is not equal to the sum of the monthly inflation rates because it is not computed as the difference in logged values of the index across periods. It is calculated as the difference in the index level between both periods over the initial period. This approach is consistent to the way the Argentine Statistical Office (INDEC) computes both monthly and annual inflation rates. See the Appendix for details.

<sup>9</sup>Santiago has 55%, Buenos Aires 32%, Caracas and Bogota 22%, and Rio de Janeiro 8% of the total population in each country).



which is studied in detail in the next section.<sup>10</sup> More importantly, they also have similar CPI methodologies. All indexes are weighted using official CPI weights, with the exception of Venezuela, where the online data cannot be classified into the standard official categories. In that case, an un-weighted index is constructed with a simple geometric average of all price changes observed each day.<sup>11</sup>

#### *4.1. Online and Official Data*

The online and official supermarket indexes are shown in Figures 1 and 2. Table 2 shows the summary statistics in each country. The “official supermarket” index is a weighted average of the “Food and Beverage” and “Household Products” official indexes. These are the sub-component of the CPI that are directly comparable to the online price indexes built in this paper. Using the CPI as the basis for comparison would not change the results, because the supermarket index is closely related to the CPI in these countries, as shown in Table 2.<sup>12</sup>

The graphs in Figure 1 show a remarkable ability of online indexes to track the main inflation trends over long periods of time. Although there are periods when online inflation is rising faster (or slower) than official estimates, over time both indexes follow a common trend. Indeed, Table 2 shows that the average annual inflation rates for the period are nearly identical in the online and official samples for all countries. This happens both in a low-inflation country like Chile, where the annual rate is 3% online and 3.19% officially, and also in a high-inflation country like Venezuela, where the annual rate is 27.43% online and 29.38% in official data.

---

<sup>10</sup>According to the 2009 WDI (World Bank), the internet user penetration is 30.4% in Argentina, 39.2% in Brazil, 33.9% in Chile, 45.5% in Colombia, and 31.2% in Venezuela.

<sup>11</sup>The unweighted index implicitly weighs products by the number of items within each sub-category. For example, milk comes in many forms and product sizes, so when all individual products are included in the index, “milk” will receive a relatively large weight in the supermarket index. Unweighted indexes can provide a good approximation if there is a link between the variety of products and the share of expenditure that a category represents. Consistent with this idea, [12] find a strong correlation between the number of SKUs and purchase volume in a large US retailer. The results of this paper hold if unweighted indexes are used, as shown in the Appendix.

<sup>12</sup>See the Appendix for more comparisons between the all-items CPI and the official supermarket index

In the annual inflation series of Figure 2, Chile and Colombia stand out for the close match in both the level and dynamics of annual inflation rates. In Chile, the online and official annual estimates are nearly identical over time, with inflation peaking in April 2009 at 8.7%, falling to -1.3% in February 2010, and climbing back to 3.5% in March 2011. In Colombia, online and official annual inflation series also show very similar patterns over time, with inflation dropping from about 8% in late 2008 to about 3% in late 2009, and remaining relatively stable ever since. In Brazil and Venezuela, the online data can also approximate the average level of annual inflation, as seen in Table 2, although the dynamics of annual inflation are less synchronized. This is evident in both the graphs and the low correlations of annual series shown in Table 2. For Brazil, online and official inflation alternate periods of higher inflation (the correlation improves after June 2010). In Venezuela, the online index puts annual inflation in the 25% - 35% range during this period, consistent with official numbers, but it peaked in December 2009, six months before the official index. This may be reflecting a delay of price adjustments in “public” supermarkets owned by the government, which were 25% of the CPI sample in 2008.

The differences in the online-official matching across countries appears to be linked to the representativeness of the retailer used in online data, shown in Table 1. In Chile, for example, the supermarket sampled has a relatively large market share of 27%, compared to only 15% for the Brazilian retailer. At the same time, the city of Santiago (where the products priced online are delivered) represents 55% of the national CPI (it was 100% until 2009), while Rio de Janeiro is only 17.3% of the CPI sample in Brazil.

In all countries, as we move from annual to monthly rates, the matching between online and official series becomes weaker. The correlation between series shown in Table 2 tends to be lowest for the monthly rates. This is linked to the use of data from a single retailer in each country, which in the short run may be able to adjust its prices slower or faster than the economy as a whole. Indeed, many of the temporary deviations from the official index in Figure 1 appear to be driven by idiosyncratic characteristics of these retailers. For example,

the retailer in Brazil tends to increase prices in “waves” rather than continually over time. Also, the supermarket in Colombia has store-wide sales every few months, which causes the index to drop temporarily for about a week. Still, these deviations tend to be corrected after a few months, improving the matching over longer periods of time.

In short, these results show that online price indexes –even when constructed with data from a single retailer– are able to match both the average level of annual inflation and the long-run inflation trends in these countries. Their ability to match high-frequency inflation dynamics appears to depend greatly on the representativeness of the retailer (market share) and the CPI relevance of the city where the online products are collected.

## **5. Argentina**

This section focuses on Argentina, where the official estimates of inflation have become widely discredited in recent years. It first documents large differences between official inflation estimates and those obtained independently with online data. It then provides several alternative indexes to evaluate the robustness of these findings, and concludes that the best approximation to the official numbers is simply to use one-third of the actual inflation rate observed online.

### *5.1. Government Intervention in the Statistical Office*

Argentina’s inflation grew steadily after 2003 as a result of an expansionary monetary policy designed to stimulate consumption and avoid the appreciation of the currency. Inflation became a politically-sensitive issue in 2006, when the annual inflation rate increased over 12%. A combination of subsidies and price controls failed to contain prices, so in January 2007 the government made a drastic decision: to take direct control of the National Statistics and Census Institute (INDEC) and fire the people responsible for computing and publishing the CPI. Since then, official statistics have become widely discredited in the media

and academic circles.<sup>13</sup> INDEC's former employees denounced a manipulation of the data, claiming that the Secretary of Commerce had demanded access to the secret list of retailers being sampled. During 2007, official estimates of annual inflation remained below 9.7%, while surveys of inflation expectations conducted by Di Tella University reached 30% by late 2007. In April 2008, the INDEC announced methodological changes that included a small adjustment in category weights, a reduction of the number and type of goods sampled, and the decision to stop using data collected by provincial statistical offices.

Gradually, economists and private institutions started to monitor the prices for small baskets of products and reporting inflation rates that were significantly higher than official estimates.<sup>14</sup> Provincial governments also computed regional inflation estimates inconsistent with INDEC's estimates. The government has repeatedly claimed that these metrics are flawed because the samples are small and not representative. Over time, it has increased its pressure on economists and institutions publishing alternative inflation estimates. In February 2011, several economists received official letters threatening them with large fines and jail-time if they continued to publish their own inflation results.

In the current context, online data provides a unique opportunity to measure alternative inflation rates in Argentina. The government cannot interfere with the data collection, which is done remotely, the sample size is several order of magnitude larger than in other private estimates (even than the CPI itself), and perhaps more importantly, Section 4 shows that the same online data and methodologies provide close approximations to official inflation rates in other Latin American countries.

## *5.2. Differences with Official Data*

The online supermarket index for Argentina, built using online data and official category weights, is compared to an equivalent official index in Figure 3.

---

<sup>13</sup>See [1], [7], and [11].

<sup>14</sup>I was one of them. Using only 150 items from a couple of online supermarkets, I created a Webster ([www.inflacionverdadera.com](http://www.inflacionverdadera.com)) that published daily inflation statistics. The CBA Index shown in Figure 5 was part of that project.

In Argentina's case, the online price index follows a completely different trend. The accumulated difference with official data has continued to grow steadily over time. Between October 2007 and March 2011, the online index increased over 100%, while the official index grew only 35%.

The annual inflation series in Figure 2(b) present two main results. First, online inflation has been consistently between 2 to 3 times higher than official inflation rates. Second, the online and official estimates share a surprisingly similar pattern over time: inflation fell in early 2009, increased in early 2010 and fell slightly in early 2011. In fact, as seen in Table 2, the correlation in annual inflation rates is quite high in Argentina, just like in Chile and Colombia.

The monthly inflation rates tell a similar story. The online rate has been consistently above the official rate for the whole period. A simple OLS Regression between monthly online and official rates, shown in Table 2, suggests that the official index is missing an average of 84 basis points every month compared to the online inflation. In fact, Argentina is the only country where this regression yield a statistically significant constant. Table 2 also shows how the official monthly inflation rate is far more stable in Argentina than in any of the other country (relative to the observed online volatility).

These results are puzzling because they are exactly the opposite to the ones observed in the other countries. Online data cannot match the official *level* of inflation in Argentina, which on average is 20.14% vs the official 8.38%, but it does an excellent job in matching even the short-run dynamic behavior of inflation rates. Indeed, the contemporaneous correlation in the monthly inflation series is higher in Argentina than in any other country. If the data from this single retailer were not representative, as it appears to happen in the case of Brazil, then we would expect the short-term dynamics to be affected first.

### 5.3. Robustness

This section considers several robustness tests to find alternative reasons for the differences in Argentina's online and official inflation estimates.

### 5.3.1. *Alternative Supermarket*

The fact that online data can match official inflation rates so well in other countries, but so badly in Argentina, suggests that the differences are not caused by general characteristics of online data. Still, there may be concerns that the online data used in Argentina are not as good or representative as the data used in other countries.

One possibility is that online prices behave differently than offline prices in this particular retailer. However, in another paper, [3], I directly compared the online and offline prices of a small sample of goods in this supermarket, and found that online and offline prices had a similar levels of inflation in January 2009. Even if that changed, it seems highly unlikely that online and offline prices can behave so differently for such a long period of time. For example, assuming that online and offline prices were the same in October 2010, by March 2011 online consumers would have been paying over 60% more than offline buyers in the same supermarket.

Another possibility is that this particular retailer is not representative of the country as a whole. To test this, I construct an online price index using data from Retailer #2. As described in Section 2, this is a supermarket with widely different characteristics: a small, online-only supermarket that targets high-income people. There is no category information for products in this retailer, so the index includes only a simple geometric average of daily price changes among all products in the retailer each day. This is the same methodology used for Venezuela, where it does a relatively good job at matching official inflation.

The online indexes from both alternative retailers are shown in Figure 4. Supermarket # 2 had slightly *higher* inflation before the recession in mid-2009, and *lower* inflation after that. In terms of the average level of inflation, however, the results across supermarkets are quite similar: inflation is much higher than reported in official statistics.

### 5.3.2. *Alternative Data: Price Controls*

Another alternative explanation for the differences with official data is that the INDEC may be using prices for goods under price controls. These are identified online with a sign

that reads “Government Agreement Price”, which was coded into an indicator variable in the database.

Price controls for a given set of goods tended to last only a few weeks during this period, so there are two possible outcomes: these products could have lower inflation rates overall if the firms are concerned about attracting too much attention (and more controls later on), or alternatively, these products could have higher inflation rates to compensate for the price-adjustments that could not be implemented while the government imposed the restrictions. Figure 4 suggests the latter happened. The inflation rate for price-controlled items increased above 40% in 2010.<sup>15</sup>

### *5.3.3. Alternative Method: The Basic Food (CBA) Index*

Beyond the data, it could be that the methodology of the supermarket index, which works so well in other countries, is somehow incompatible with official CPI statistics in Argentina. So, instead of constructing a supermarket index, I now focus my comparisons on a much simpler indicator: the “Canasta Básica de Alimentos” (CBA). This is a Basic Food Index used to measure the level of extreme poverty in the country. Its methodology is well documented, with details of the exact weighing and characteristics of the products that underlie its construction, so I can replicate the exact same basket of goods. This basket has a much smaller number of products within each category, instead of every item in the store.<sup>16</sup>

Figure 4(a) compares the online and official CBA indexes. Once again, the differences are huge and persistent over time: by March 2011, the online CBA index had an accumulated inflation of 91%, while the official INDEC CBA had increased by only 31%. Interestingly, there is a spike in online prices during March 2008 which perfectly coincides with the timing of the massive “Farm Strike” that took place in the country. This was a strike of farmers in response to the attempt to introduce a variable-scale export tax regime on several key

---

<sup>15</sup>Causality is not clear from this exercise. Maybe the price limits were imposed precisely because these have traditionally been high-inflation products that the government wanted to control.

<sup>16</sup>Details for the basket and weights is provided in the Appendix.

crops, which in practice meant an increasing the tax rate to 45% for soybeans (Argentina's main agricultural export). Farmers organized road blocks from March to June and caused severe shortages of basic food products for several weeks. The most dramatic price increases occurred during the first strike, which started on March 13th, 2008 and lasted until April 3rd, 2008. The effects were mostly temporary, with the online index going back to its original trend by June 2008. By contrast, the official CBA index completely ignored these shocks.<sup>17</sup>

#### *5.4. Best Approximation: 1/3 rd of the observed inflation rate*

The results so far strongly suggest that the INDEC is manipulating the official inflation estimates. How this is done is an open question, but the main results in Figure 3 suggest a simple answer: that the government is reporting just a fraction of the real inflation rate, usually between a third or one-half of the actual numbers. In fact, a simple OLS regression of the official and the online annual inflation series yields a coefficient of 0.36. To test this hypothesis, we can run a simple simulation: an index built using only a third of the inflation rate observed each day.

This simple approach yields an index with a remarkable resemblance to the official data, as seen in Figure 6. This suggests that the way the data is being altered is far simpler than commonly assumed. After all, the INDEC is a large organization, with many employees involved with the data collection and construction of the price indexes. Instead of altering the item-level prices or the intermediate indexes, it must be easier for the government to change the aggregate results, which are computed by just a handful of people at the end of the CPI calculation. <sup>18</sup>

---

<sup>17</sup>The online supermarket index also has a sharp increase in march 2008, but the effects appear to be permanent in this index. It is possible that supermarkets used this shock as an opportunity to justify increases in the aggregate index that had been delayed for some time. In particular, the year before the strikes the government was very aggressive with the implementation of price controls and boycotts against supermarkets that raised prices. The farm strikes, which should have had only a temporary effect on food, could have provided a way to hide permanent increases in other goods as well.

<sup>18</sup>Several other alternatives methods are considered in the Appendix. Some are methodological, such as the cell-relative imputation of missing prices or the use of unweighted indexes, and others use special subsets of the data, such as the lowest inflation goods in each sub-category. None of them is able to match the low inflation estimates present in the official data.



### 5.5. *Implications for Other Statistics*

Biases in inflation raise suspicions in other statistics, like poverty and Real GDP. For poverty measurements, the implications for the difference with the official data in the CBA index shown in Section 5.3.3 is straightforward. Every quarter, INDEC uses the cost of the CBA basket to see how many individuals are in extreme poverty conditions. Taking the cost of the official basket in the first quarter of 2008, and adjusting it with the CBA inflation rate observed online, the basket in July 2011 would have a cost of \$259.5 Argentine pesos. Using INDEC's own income statistics, this implies that 6.69% of the population was under extreme poverty, compared to only 2.5% reported in official statistics. Adjusting the CBA to obtain the broader "Canasta Basica Total" (CBT), which adds non-food items to the basket, the level of poverty becomes 25.9% compared to the 9.9% officially reported.<sup>19</sup>

For real GDP estimates, we can look at how CPI and the GDP deflator have behaved in the past decade. Figure 7 plots both series from 1994 to 2011. The data from 1994 and 2006 shows that both series were closely correlated, which is the expected behavior under normal conditions.<sup>20</sup> However, since 2006, the GDP deflator has increased significantly faster than the CPI. This means that government has recognized much higher price increases in the GDP deflator. This is not surprising. If it had increased as little as the CPI, the real GDP growth rate would have been in double digits for several years, which would be impossible for the government to justify. Furthermore, allowing for a higher GDP Deflator has little political cost, as it is rarely followed by the media or the general public as a metric for inflation.

Still, the GDP Deflator has increased considerably less than the online index since 2007. To get a lower bound for the real growth rate, an "Adjusted Real GDP" can be calculated using the online index to deflate the official Nominal GDP. The results, shown in Figure 7,

---

<sup>19</sup>The CBA cost is multiplied by 2.16 to obtain the CBT cost. This coefficient, reported by the INDEC, is the inverse of the share of food over all basic household expenditures.

<sup>20</sup>There are reasons to expect small and temporary differences between the GDP deflator and the CPI. The CPI includes only consumer prices, while the GDP deflator includes the prices of goods purchased by the government and firms. In addition, the CPI includes prices of both imports and exports, while the GDP deflator focuses exclusively on domestically produced goods. These reasons, however, do not seem able to justify the persistent differences observed since 2006.

provide GDP growth rates of -4% in June 2009 (right before the government lost the parliamentary elections), 5% in June 2010 , and only 0.5% in March 2011. Of course, these numbers should be taken as a rough approximation, because an alternative GDP cannot be estimated directly with online data. Still, they provide a lower-bound estimate for the true growth performance of the country.

## 6. Conclusions

Online price indexes, constructed using a combination of online data and official methods and weights, are capable of matching both the level and main dynamics of official inflation in Brazil, Chile, Colombia, and Venezuela. The matching is best at annual frequencies and improves when the data comes from supermarkets with large market shares and cities that are more representative of the country as a whole.

The results for Argentina confirm the suspicion that the government is manipulating the official inflation rates. It is the only country where online inflation deviates significantly from official estimates over time. Two things are surprising: the magnitude of the difference, and its persistence over time. Indeed, online inflation has been consistently between two and three times larger than in official estimates for over 3.5 years. On average, the online index had an annual inflation rate of 20.14%, compared to just 8.38% in official data. The accumulated difference was 65% by March 2011. Surprisingly, the difference lies exclusively in the *level* of inflation reported every period, not the dynamic behavior of inflation rates over time.

Several robustness exercises were considered, but none seems able to account for these large discrepancies with official data. The best approximation to the official series is to simply assume that the inflation rate is one-third of the one observed in the online data.

There is no obvious reason for why the government chose to manipulate the price index. Some economists believe it was to save on interest payments for inflation-indexed bonds, but this would have been a very short-sighted strategy, considering the amount of uncertainty

this has introduced into the economy. It is likely that the government was just trying to hide what it thought was a temporary rise in inflation, and as time went by, it became increasingly harder to recognize the mistake. In any case, after nearly four years of manipulation, the government has shown little interest in making any changes to the way it computes and reports the inflation statistics. Independent economists are now being threatened with fines or jail time if they continue to publish alternative estimates. A collaboration announced with the IMF in 2010 to construct a new price index appears to have been just a way to delay any sanctions. Some policies are starting to show inconsistencies with the official data. For example, the government allowed unions to request salary increases of 25% in 2011, three times higher than the official inflation rates. Sooner or later, the official estimates will have to become accurate. In the meantime, online indexes can provide a good approximation to the real inflation rate in the country.

- [1] **Barrionuevo, Alexei**, “Inflation, an Old Scourge, Plagues Argentina Again,” *New York Times*, February 2011.
- [2] **BLS**, “Handbook of Methods,” in “BLS Handbook of Methods” Chapter 17 2009.
- [3] **Cavallo, Alberto**, “Scraped Data and Sticky Prices,” *Working Paper*, 2010.
- [4] **Chile, INE**, *Manual Metodológico del Índice de Precios al Consumidor IPC Nacional - Base 2009* 2009.
- [5] **Colombia, DANE**, *Metodología del Índice de Precios al Consumidor* Colección Documentos, 62 ed. 2009.
- [6] **de Chile, Banco Central**, *Índice de Precios al Consumidor* February 2010.
- [7] **Economist, The**, “Lies and Argentine Statistics,” *The Economist*, April 2011.
- [8] **Greenless, John and Robert McClelland**, “Addressing Misconceptions about the Consumer Price Index,” *BLS Monthly Labor Review*, 2008, *August*.
- [9] **IBGE**, *Sistema Nacional de Índices de Preços ao Consumidor*, Vol. 14 of 5, Instituto Brasileiro de Geografia e Estatística, 2007.
- [10] **INDEC**, *Índice de Precios al Consumidor Gran Buenos Aires* Documento Metodológico Nro 13, Instituto Nacional de Estadísticas y Censo, 2001.
- [11] **Jueguen, Francisco**, “Entrevista con Graciela Bevacqua,” *La Nación*, July 2009.
- [12] **Simester, Duncan and Eric Anderson**, “Menu Cost and Price Rigidities: Micro Evidence,” *Working Paper*, 2011.
- [13] **Venezuela, INE**, *Índice Nacional de Precios al Consumidor* 2008.

# Figures

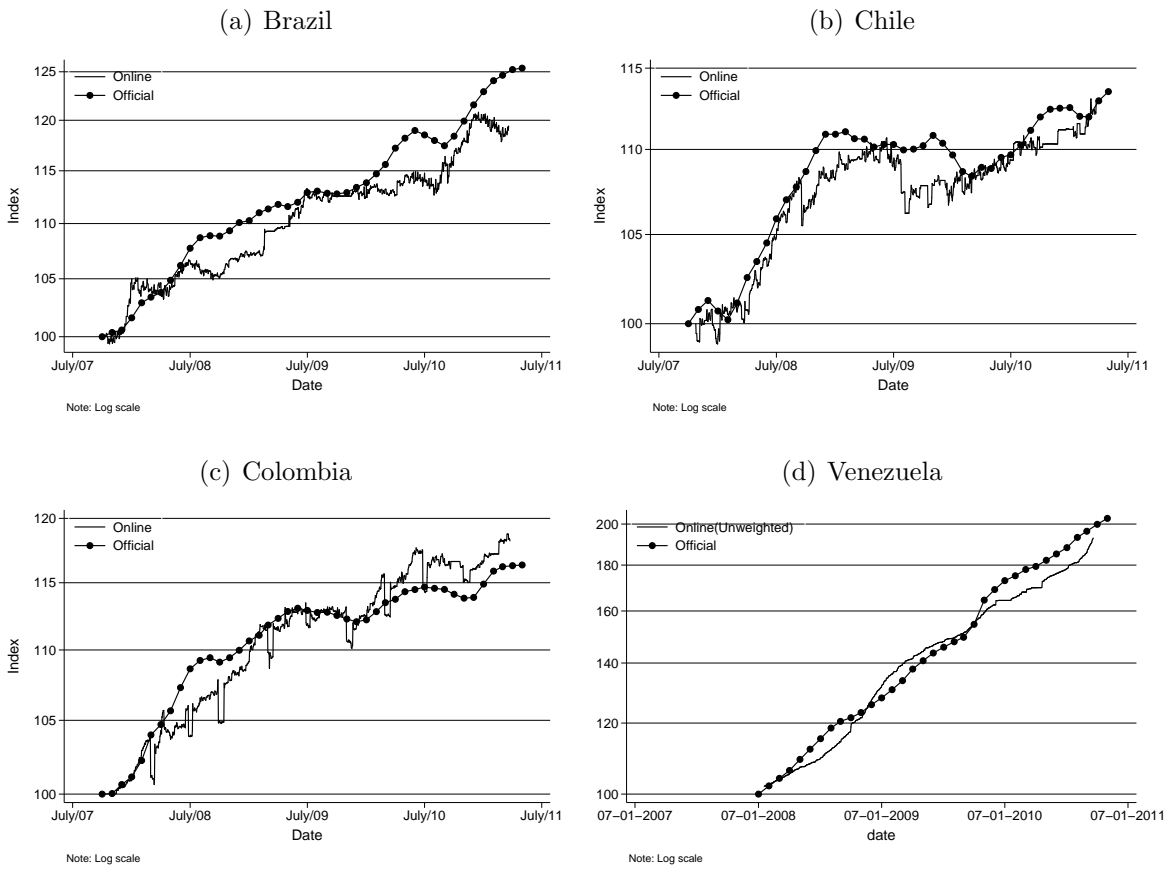


Figure 1: Online and Official Indexes in Four Latin American Countries

Notes: The daily online supermarket index is constructed with a online price information and official category weights. The only exception is Venezuela, whose retailer has no category information and therefore the online index is built with a simple geometric average of all price changes observed each day. The official index is an equivalent supermarket indicator, constructed as a weighted average of the Food and Beverage and Household Product official price indexes in each country.

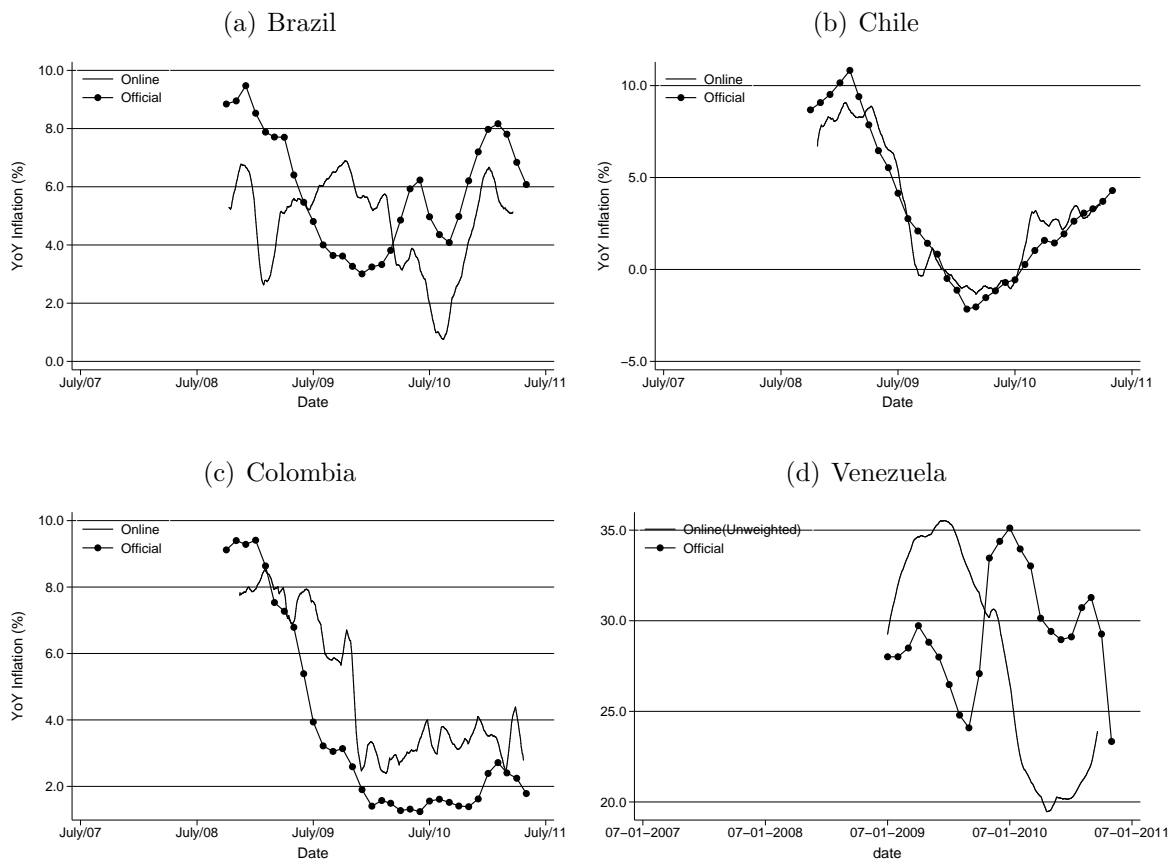


Figure 2: Online and Official Indexes - Annual Inflation Rate

Notes: The annual inflation rate is the percentage change in the average of the index during the past 30 days with respect to the average of the index in the same period a year ago.

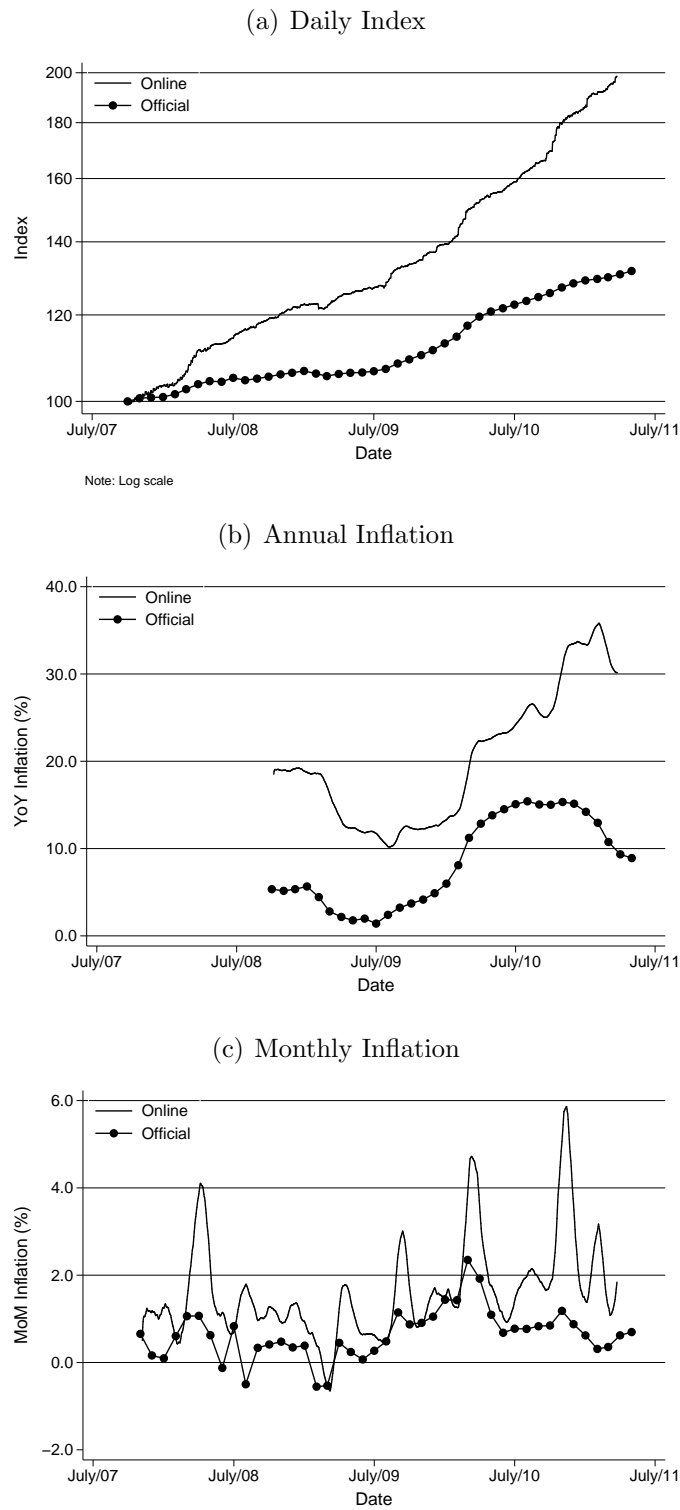
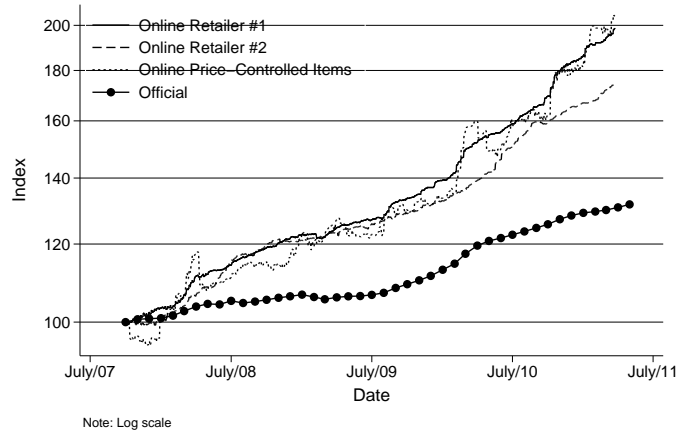


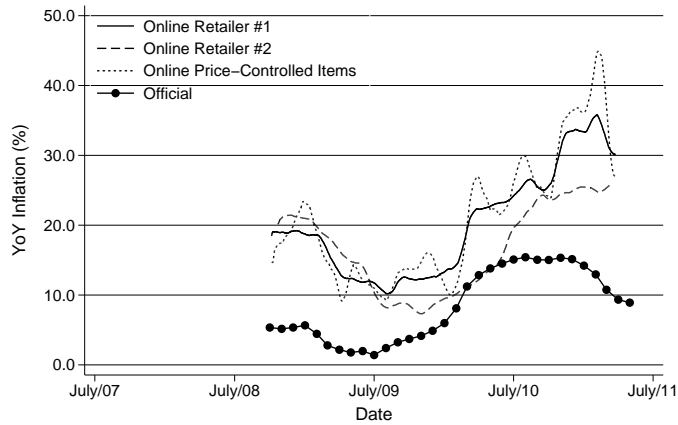
Figure 3: Online Supermarket Index in Argentina

Notes: The daily online supermarket index is constructed with a online price information and official category weights. The monthly official index is an equivalent supermarket indicator, constructed as a weighted average of the Food and Beverage and Household Product official price indexes. The annual inflation rate is the percentage change in the average of the index during the past 30 days with respect to the average of the index in the same period a year ago. The monthly inflation rate is the percentage change in the average of the index during the past 30 days with respect to the average of the index in the previous 30 days.

(a) Daily Index



(b) Annual Inflation



(c) Monthly Inflation

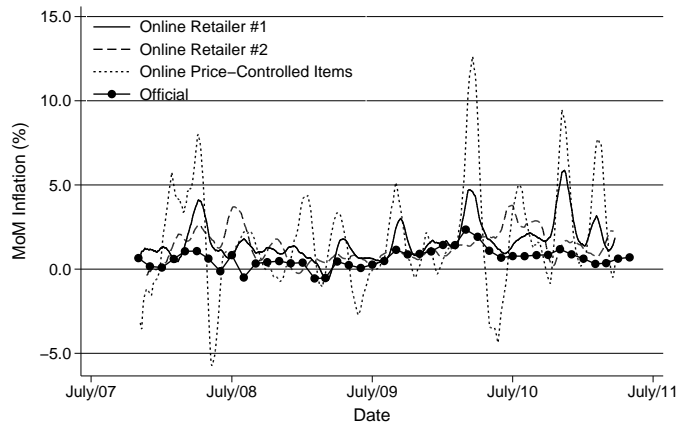
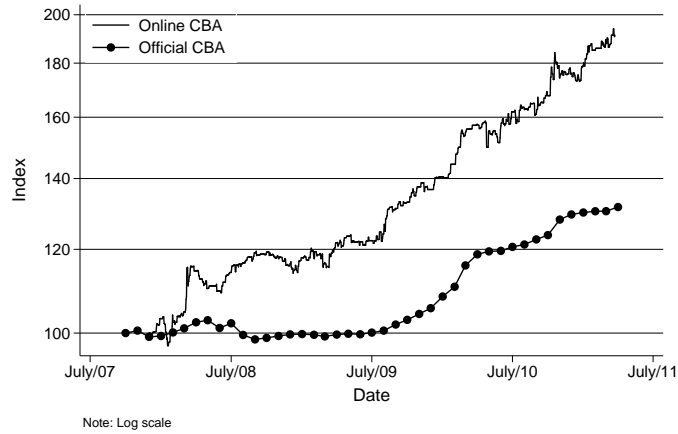


Figure 4: Different Supermarkets and Price Controls

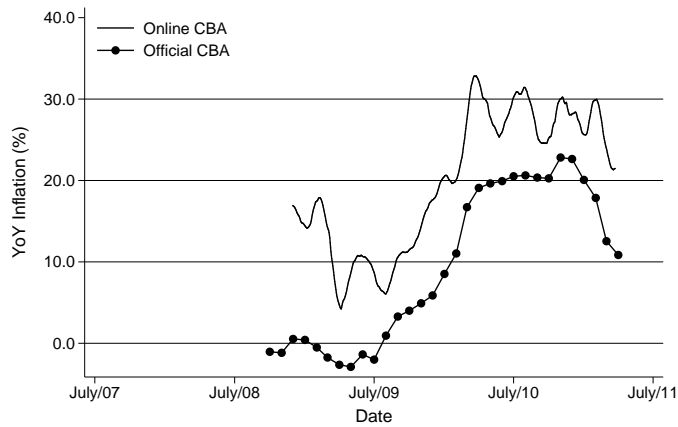
Notes: The Price-Controlled Index uses 597 products that were under a price control imposed by the government at any point in time during the sample period. These items can be identified by an image in the retailer's Webster. Argentina's Retailer #1 is the largest supermarket chain in the country. Retailer #2 is a smaller supermarket that sells exclusively online to a higher-income population. The index for Retailer #2 is an un-weighted index because the data lacks information on product categories.



(a) Daily Index



(b) Annual Inflation



(c) Monthly Inflation

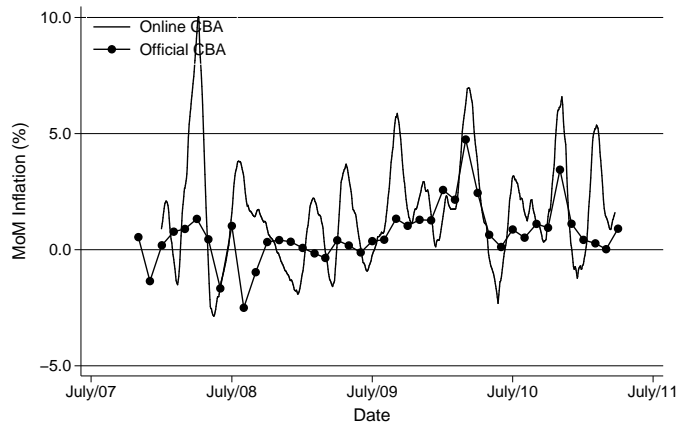
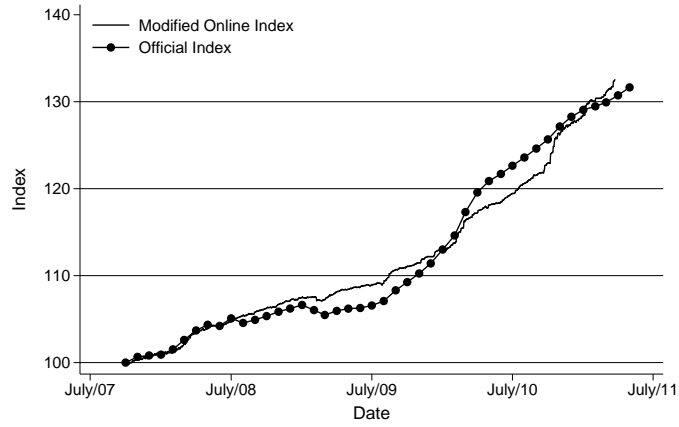


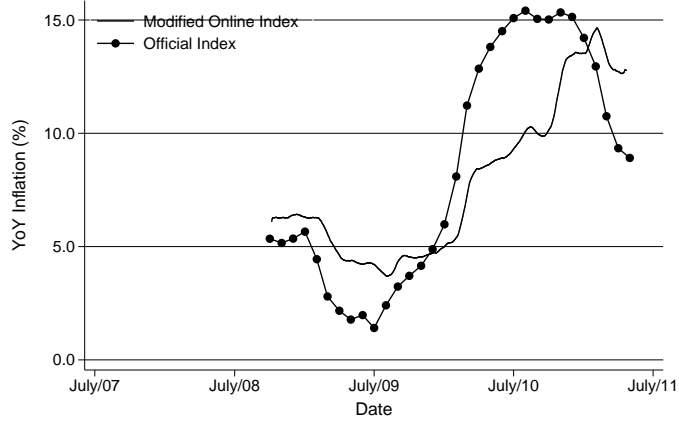
Figure 5: Subsistence Food (CBA) Index

Notes: The “Canasta Basica de Alimentos” (CBA) index is a subsistence food indicator used to measure the level of extreme poverty in the country. It uses 45 carefully selected goods, with official weights and one price from each one of the two retailers available in the country’s dataset. More details are provided in the Appendix.

(a) Daily Index



(b) Annual Inflation



(c) Monthly Inflation

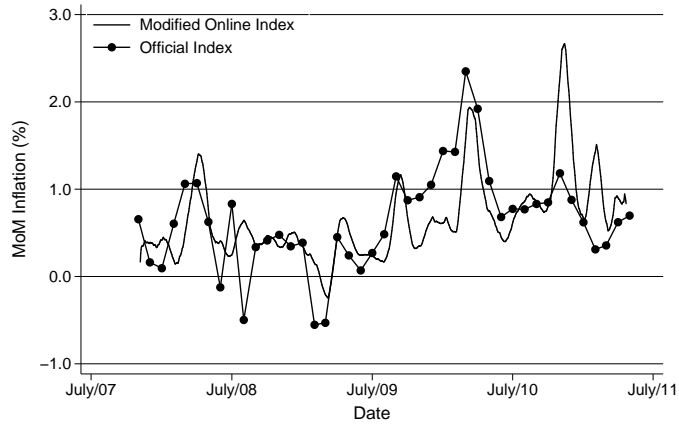
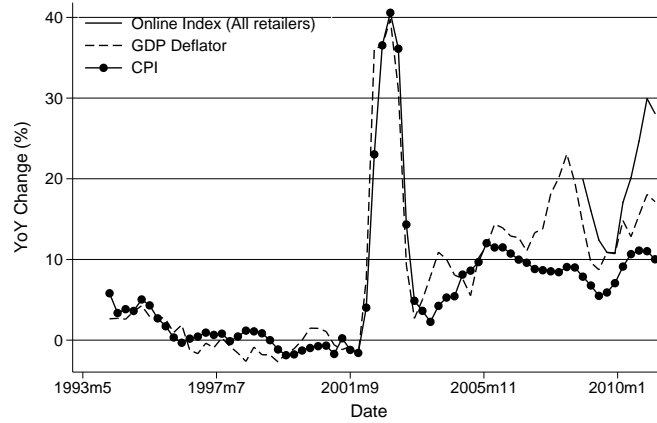


Figure 6: Approximation:  $1/3$  of the observed inflation rate

Notes: The *modified* online index uses only a third of the daily inflation rate observed with online data.

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



(b) Real GDP - Annual Growth Rate

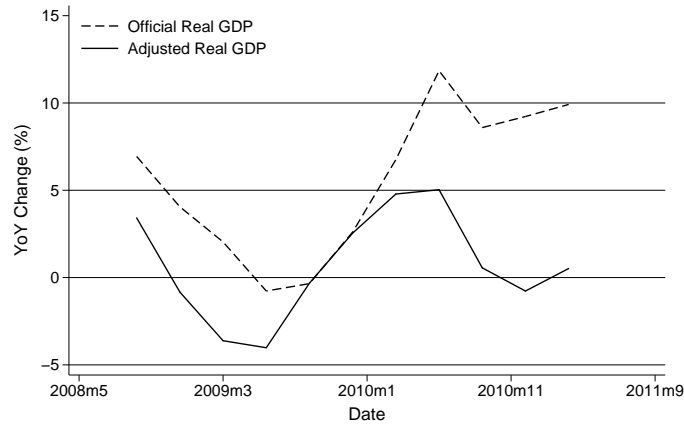


Figure 7: Implications for Real GDP Growth

Notes: The Online Index in (a) is an average of the supermarket indexes of retailers #1 and #2. The adjusted real GDP series are computed by multiplying the official nominal GDP with an adjusted GDP deflator. The adjusted GDP deflator assumes that the official deflator in Q4 of 2007 increased at the same quarterly rate as the Online Index until Q1 2011.

## Tables

Table 1: Online 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
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
Category Indicator	Yes	No	Yes	Yes	Yes	No
CPI Weights Covered	48.51%	-	27.93%	31.00%	28.44%	-
Retailer Market Share**	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 4. \*\*Market shares are based on claims posted on the corporate webpages of each supermarket's parent companies.

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

Note: The top panel shows that Argentina is the only country where online data does not approximate the average official annual inflation rates. The third panel shows that the correlation between the monthly inflation rates is higher in Argentina than in any of the other countries. This shows that discrepancy is in the level of inflation reported, not its dynamic behavior over time. The third panel reinforces this idea with a simple OLS regression of the official monthly rate and 12 lags of the online monthly rate. The R2 is highest in Argentina, which is also the only country where the constant is statistically significant. The fourth panel shows that the official monthly inflation rate is indeed surprisingly stable in Argentina compared to both the online index and the volatility observed in a high-inflation country like Venezuela.

# Online and Official Price Indexes: Measuring Argentina's Inflation Alberto Cavallo

MIT Sloan  
June 17, 2011

## Appendix A. Monthly and Annual Inflation using Daily Data

Formally, the monthly scraped inflation  $\pi_t^m$  for month  $m$  is defined as the percentage change in the index in the past 30 days versus the previous 30 days,

$$\pi_t^m = 100 * \left[ \frac{\frac{1}{30} \sum_{x=0}^{29} I_{t-x}}{\frac{1}{30} \sum_{x=30}^{59} I_{t-x}} - 1 \right] \quad (\text{A.1})$$

On the last day of each calendar month,  $\pi_t^m$  is comparable to the monthly inflation reported in official statistics.

Similarly, annual inflation is computed as the percentage change in the index in the past 30 days over the same period 365 days ago,

$$\pi_t^y = 100 * \left[ \frac{\frac{1}{30} \sum_{x=0}^{29} I_{t-x}}{\frac{1}{30} \sum_{x=365}^{394} I_{t-x}} - 1 \right] \quad (\text{A.2})$$

On the last day of each calendar year,  $\pi_t^y$  is comparable to the annual inflation rate reported in official statistics.

## Appendix B. More Robustness Tests

### *Appendix B.1. Cell-Relative Imputation and unweighted Index*

I consider two more robustness exercises related to the methodology.

The first exercise uses a different approach to impute missing values within price series. The method used in previous sections was to fill missing prices with the last available price for each product. I do this because the gaps in my data last only a few days, but this is not how most statistical offices deal with missing values in their monthly series. The standard approach, also used in Argentina, is to impute missing prices using the average price change of similar products, whose prices are not missing. Methods vary slightly across countries, but I follow the “cell-relative” approach used by the Bureau of Labor Statistics (BLS): if a product is missing on a particular day, I do not use that product for the calculation of that day’s inflation, but I impute a price for it equal to the previous price times the average price change for products in the same category that day. In theory, the two different methods should not impact long-term inflation estimates, but could yield differences in some of the short-run dynamics. Indeed, Figure B.1 shows that the monthly inflation estimates are nearly identical for most of the sample, with the exception of March and November 2010. The overall differences are small, and not nearly strong enough to explain any of the differences between online and official estimates.

The second robustness test is to use an unweighted index, just as the one constructed for Venezuela. This index is computed with a geometric mean applied to all price changes observed each day in the store, without any categories or relative weights. Naturally, the effect of a geometric mean is to reduce the impact of products that have relatively large price changes (either up or down). It is implemented by the BLS and other statistical offices at a sub-category level in order to approximate the effect of within-category substitution (for example, when the price of a type of apple goes up, consumers will tend to shift to other apples). In this case, I am pooling together goods from different categories, effectively allowing cross-category substitutions as well. In theory, we should expect to see less inflation when it is rising, and more inflation when it is falling. That is precisely what Figure B.1 shows. In particular, this approach produces considerably less inflation in Argentina from 2010 onward, because of a lower impact of high-rising prices in March and November of that year. Still, the inflation rate is twice as high as the official estimates. So, in other words, even if INDEC were to be allowing for cross-category substitutions, the effect does not appear to be strong enough to explain their results.<sup>21</sup>

#### *Appendix B.2. Lowest-Inflation items*

Could a selection of the goods included in the index help reduce the observed inflation rate? It is possible that the government is monitoring prices that it somehow knows are going to have lower inflation rates? This section uses an extreme assumption: that the government is able to select, within narrowly defined categories, only the goods that have the lowest inflation rate over this whole period. This would not be a realistic alternative for the INDEC, given that it is hard to know ex-ante which goods will have the lowest inflation rates, but at least it can provide a lower-bound inflation rate for a strategy that favors lower-inflation brands or items.

In this data the way to identify close substitutes is to use the URL where the product is located. Within these categories, the product with the lowest inflation rate for the whole period is selected and included in the index.

As expected, Figure B.2 shows that the observed inflation is much closer to official estimates, particularly in the first semester of 2010. However, the annual inflation rate is often twice as high as the official estimates reported. There is no subset of the data, however arbitrarily selected, that can yield the low rates of inflation reported by the government.

---

<sup>21</sup>This exercise also suggest that the unweighted index for Retailer #2, presented in Section 5.3.2, might be underestimating inflation in the same period. This would move it even closer to the results for Retailer #1, and even further away from the official estimates.

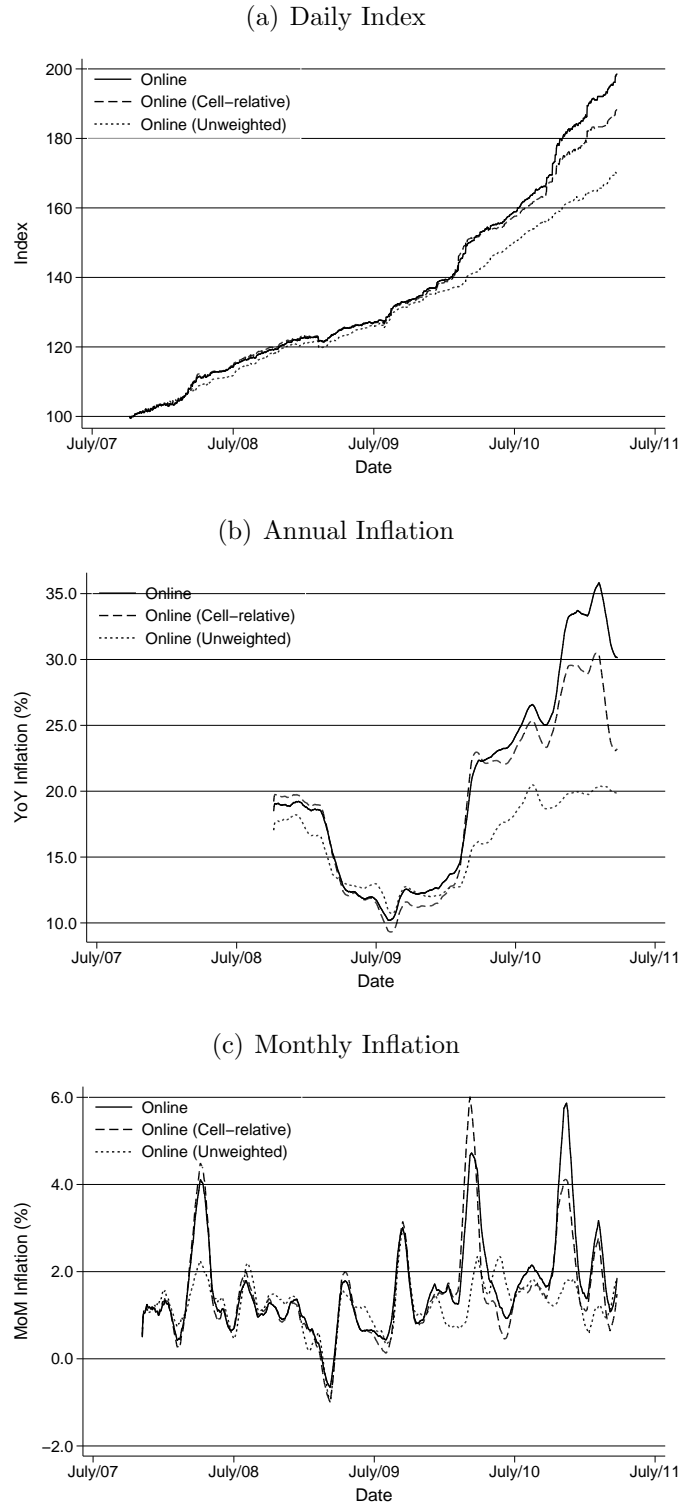
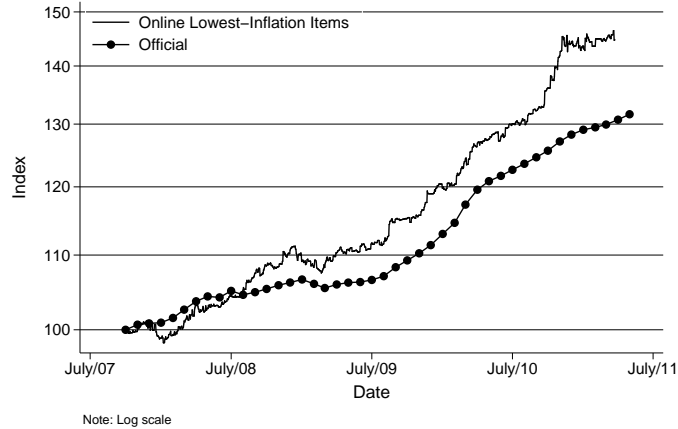


Figure B.1: Robustness: Cell-Relative Imputation

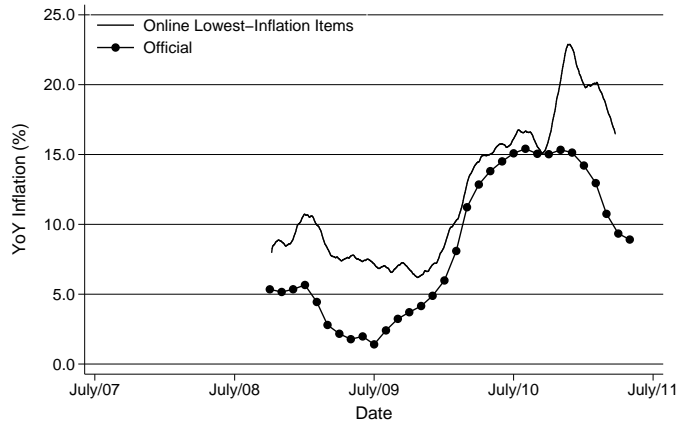
Notes: The “cell-relative” index uses the BLS method to impute missing prices within price series, using the average price change of similar non-missing products in the category. The un-weighted index in Argentina uses the data from Retailer #1, without any category weights. It is a simple geometric average of all price changes in a given day.



(a) Daily Index



(b) Annual Inflation



(c) Monthly Inflation

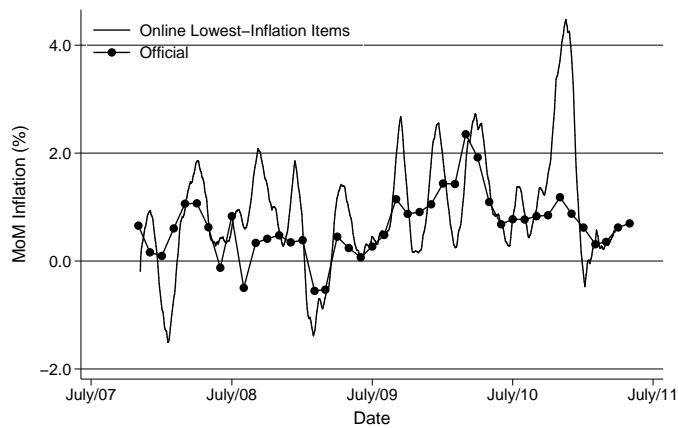


Figure B.2: Lowest-Inflation Items within URLs

Notes: This index includes only 893 products that had the lowest inflation over the whole sample period within each URL. The URL is the lowest level of aggregation available, grouping together close substitutes like "Olive Oil" or "Milk". There is only one item per URL. To select them, I first filtered products that had at least 2 years of data, then computed their total inflation rates from the first day they appear in the sample until their last day, and finally I chose the lowest-inflation item within each URL to construct the index.

Appendix B.3. Additional Tables and Figures

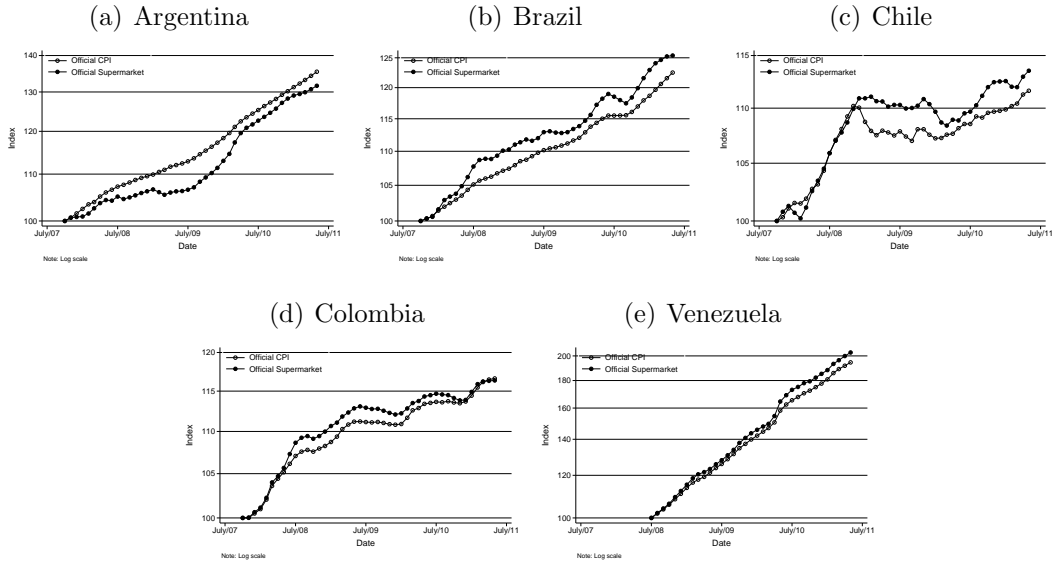


Figure B.3: Official CPI and Official Supermarket Indexes

Notes: The official supermarket index is constructed as a weighted average of the Food and Beverage and Household Product official price indexes in each country.

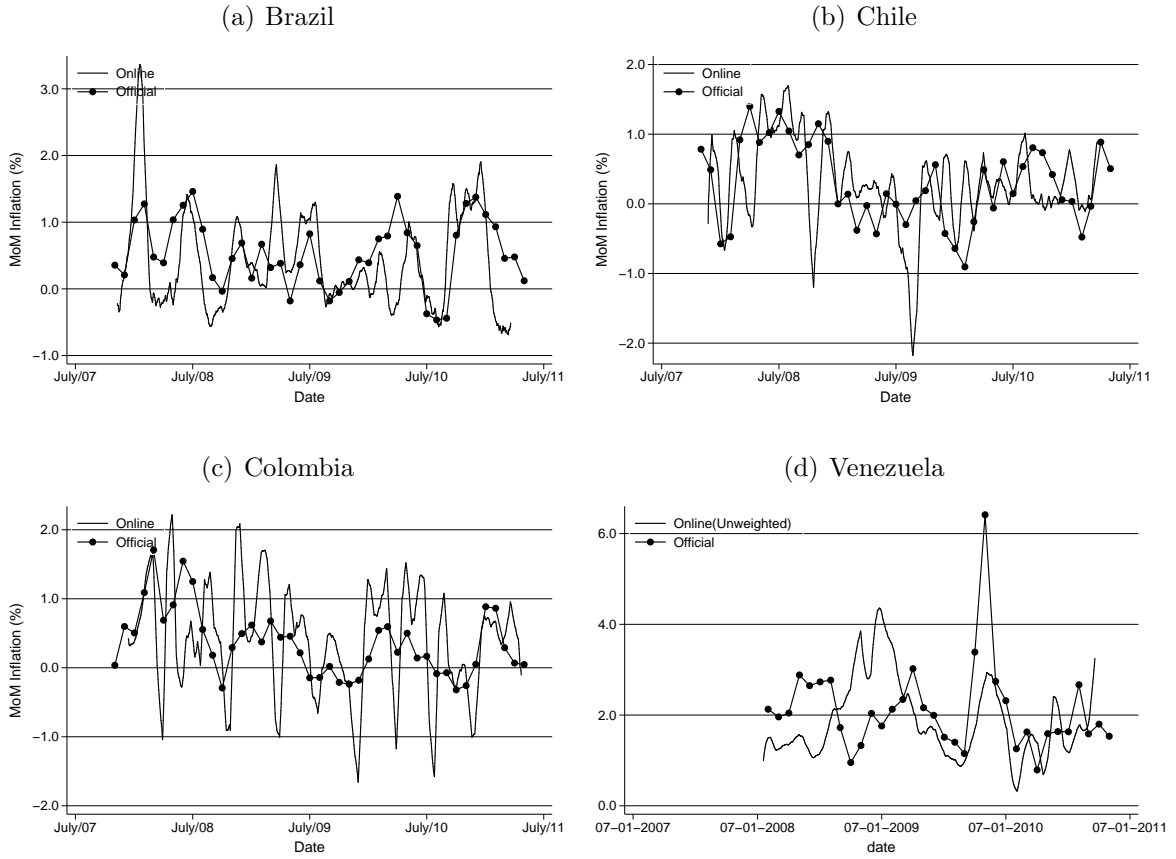


Figure B.4: Online and Official Indexes - Monthly Inflation Rate

Notes: The monthly inflation rate is the percentage change in the average of the index during the past 30 days with respect to the average of the index in the previous 30 days.

### Appendix C. The CBA index

The CBA basket is constructed with a given number of grams per product. There are 45 products included in the index, detailed in Table C.2. These products and their respective grams are set by the INDEC to meet the minimum nutritional requirements for an adult male in the 30-59 age range. I calculate the daily cost of the basket using 45 items from the scraped data that were carefully chosen to match the products in the INDEC listing. Because these items come in package sizes which do not always coincide with the number of grams in the basket, I multiplied their prices by the ratio of grams in the official listing over the actual grams in the package. These weights are shown in Table C.2. Next, I added all these weighted prices to obtain the daily cost of the basket. Finally, the index value is simply the cost of the basket at time  $t$  over the cost of the basket at the initial date  $t = 0$ .

Table B.1: 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	
Fresh Fish And Seafood	0.41	0.25	0.31	0.60
Processed Fish And Seafood	0.11	0.70	0.17	
Eggs	0.57	0.20	0.31	0.74
Milk	1.27	1.48	0.89	2.03
Cheese And Related Products	1.68	0.48	0.62	0.51
Ice Cream And Related Products		0.09	0.43	
Other Dairy And Related Products	1.47	0.24	0.36	
Apples		0.09	0.12	
Bananas		0.25	0.13	0.06
Citrus Fruits		0.21	0.11	0.06
Fresh Fruits	1.24	0.28	0.37	0.88
Potatoes		0.20	0.66	0.44
Lettuce		0.08	0.16	
Tomatoes		0.16	0.32	0.19
Other Fresh Vegetables Including Fresh Herbs	2.16	0.29	1.41	1.70
Canned Fruits And Vegetables	0.55	0.03	0.06	0.10
Frozen Fruits And Vegetables			0.07	
Other Processed Fruits And Vegetables			0.11	
Carbonated Drinks	1.79	0.70	1.48	0.82
Nonfrozen Noncarbonated Juices And Drinks	0.49	0.20	0.53	0.03
Coffee	0.17	0.36	0.15	0.31
Tea - Mate	0.49	0.02	0.16	
Other Beverage Materials	0.35			0.12
Sugar And Artificial Sweeteners	0.30	0.34	0.25	0.64
Candy And Chewing Gum	0.22	0.08	0.13	
Other Sweets		0.11	0.05	
Marmalade & Jams	0.15		0.15	
Chocolate	0.15	0.13	0.21	0.26
Butter And Margarine	0.04	0.04	0.24	
Other Fats And Oils	0.49	0.43	0.30	1.02
Soups	0.09	0.02	0.05	0.11
Frozen And Freeze Dried Prepared Foods			0.04	
Snacks				0.16
Salt And Other Seasonings And Spices	0.10	0.10	0.14	0.08
Olives, Pickles, Relishes		0.04		
Sauces And Gravies		0.18	0.24	0.25
Other Condiments	0.26	0.06		
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	
Shampoo, Bath Products	0.77	0.29	0.27	
Lamps & Lighting Fixtures			0.04	0.14
Paint, Wallpaper Tools & Supplies	2.01	1.44	0.16	
Tools	0.03		0.10	
Lawn & Garden Supplies & Insecticides	0.30		0.17	
Cleaning Products	2.06	0.42	0.60	0.64
Laundry Products		0.36	0.43	1.10
Household Paper Products	0.32	0.16	0.10	
Miscellaneous Household Products	0.32		0.28	0.37
Tires	0.23	0.18	0.05	0.04
Vehicle Parts & Equipment Other Than Tires	0.71	0.67	0.10	
Toys, Games, Hobbies, & Playground Equipment	0.52	0.48	0.42	0.31
Pet Food			0.42	
Pet Supplies		0.07	0.04	0.11
Over-The-Counter Drugs	3.40	2.93	1.79	1.42
Medical Topicals And Dressings			0.04	
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

Table C.2: CBA Index Weights

INDEC		Scraped Data		
Product (1)	Grams (2)	Product (3)	Package size in grams (4)	Index Weight (2)/(4)
Pan	6,060	Pan	500	12.12
Galletitas Saladas	420	Galletas Saladas	1000	0.42
Galletitas Dulces	720	Galletas Dulces	160	4.50
Arroz	630	Arroz Largo Fino	500	1.26
Harina De Trigo	1,020	Harina De Trigo	1000	1.02
Otras Harinas (Maz)	210	Harina Maiz	500	0.42
Fideos	1,290	Fideos	1000	1.29
Papa	7,050	Papa Negra	1000	7.05
Batata	690	Batata	1000	0.69
Azcar	1,440	Azucar	1000	1.44
Dulce De Leche	80	Dulce De Leche	250	0.32
Dulce De Batata	80	Dulce De Batata	1000	0.08
Mermeladas	80	Mermelada De Frutilla	454	0.18
Legumbres Secas				
Lentejas	80	Lentejas	500	0.16
Porotos	80	Porotos	500	0.16
Arvejas	80	Arvejas Verdes	350	0.23
Cebolla	655	Cebolla	1000	0.66
Lechuga	655	Lechuga Francesa	1000	0.66
Tomate	655	Tomate	1000	0.66
Zanahoria	655	Zanahoria	1000	0.66
Zapallo	655	Zapallo	1000	0.66
Tomate En Lata	655	Tomate Perita	400	1.64
Manzana	2,010	Manzana	1000	2.01
Naranja	2,010	Naranja De Jugo	1000	2.01
Asado	896	Asado Centro Novillito	1000	0.90
Carne Picada	896	Carne Picada	1000	0.90
Carnaza	896	Carnaza Comun	1000	0.90
Cuadril	896	Colita De Cuadril	1000	0.90
Nalga	896	Nalga	1000	0.90
Paleta	896	Paleta De Cerdo	1000	0.90
Pollo	896	Pollo Sin Piel	1000	0.90
Huevos	630	Huevo	300	2.10
Leche	7,950	Leche Entera	1000	7.95
Fresco	90	Queso Crema	200	0.45
De Rallar	90	Queso Rallado	210	0.43
Crema	90	Queso Cremoso	1000	0.09
Aceite	1,200	Aceite Mezcla	1000	1.20
Bebidas Edulcoradas	4,050	Coca Cola	2000	2.03
Bebidas Gaseosas Sin Edulcorar	3,450	Soda	2000	1.73
Sal Fina	150	Sal Fina	500	0.30
Sal Gruesa	90	Sal Gruesa	500	0.18
Vinagre	90	Vinagre Alcohol	1000	0.09
Caf	60	Cafe Molido	500	0.12
T	60	Te	50	1.20
Yerba	600	Yerba	500	1.20