BANCO CENTRAL DE RESERVA DEL PERÚ

The Distribution of the Size of Price Changes

Alberto Cavallo* and Roberto Rigobon**

* MIT Sloan.
** MIT Sloan & NBER.

DT. N° 2011-011
Serie de Documentos de Trabajo
Working Paper series
Junio 2011

The Distribution of the Size of Price Changes*

Alberto Cavallo[†] Roberto Rigobon[‡]

MIT Sloan MIT Sloan & NBER

May 27, 2011

Abstract

Different theories of price stickiness have distinct implications on the number of modes in the distribution of price changes. We formally test for the number of modes in the price change distribution of 36 supermarkets, spanning 22 countries and 5 continents. We present results for three modality tests: the two best-known tests in the statistical literature, Hartigan's Dip and Silverman's Bandwidth, and a test designed in this paper, called the *Proportional Mass* test (PM). Three main results are uncovered. First, when the traditional tests are used, unimodality is rejected in about 90 percent of the retailers. When we used the PM test, which reduces the impact of smaller modes in the distribution and can be applied to test for modality around zero percent, we still reject unimodality in two thirds of the supermarkets. Second, category-level heterogeneity can account for about half of the PM test's rejections of unimodality. Finally, a simulation of the model in Alvarez, Lippi, and Paciello (2010) shows that the data is consistent a combination of both time and state-dependent pricing behaviors.

*We are grateful to Oleksiy Kryvtsov, Dan Sichel, and seminar participants at Harvard, MIT, NBER, AEA, BLS, ITAM, Oxford, Queen's University at Belfast, and the Central Bank of Norway for helpful comments and suggestions. All remaining errors are ours.

[†]email: acavallo@mit.edu

[‡]email: rigobon@mit.edu

1 Introduction

With the availability of product-level price data used in the construction of the Consumer Price Index (CPI) from several developed countries, the micro-pricing literature in macroeconomics has become one of the most active areas of research in recent years.[1] One of the main stylized facts uncovered by this literature is that the distribution of price changes (conditional on a change) is close to a unimodal centered at zero percent, with a large share of small price changes. This finding has also been shown to hold in scanner datasets from retailers in the US.[2]

These results are important because the different theories of price stickiness have direct implications on the number of modes of the price change distribution. For example, a standard state-dependent model, such as Golosov and Lucas (2007), predicts that the distribution of price changes should have little mass near zero percent. The intuition is that small deviations from the optimal price are less costly than the price-adjustment cost, and therefore small changes are rare. In a low-inflation setting, this creates a bimodal distribution around zero percent, with both a positive and negative mode. By contrast, time-dependent models of price stickiness – such as the classical Calvo (1983) model – imply that the distribution of price changes should inherit the same properties of the distribution of cost changes, and in low-inflation setting, such costs will tend to have a unimodal shape centered around zero.[3] A third kind of model combines elements of time and state-dependent pricing, giving rise to a variety of distributions whose shape depends on the relative importance of observation and adjustment costs. Examples include Woodford (2009), Bonomo, Carvalho, and Garcia (2010), and Alvarez, Lippi, and Paciello (2010). Surprisingly, even though the modality of the distribution plays a crucial role in distinguishing the different theories of price stickiness, no paper has formally evaluated the number of and location of price-change modes.

In this paper, we study modality using three statistical tests and a new dataset that covers many countries and retailers. We go beyond the graphical analyses performed in the literature and develop a new methodology that can be scaled to test for modality in multiple countries and sectors.[4] Contrary to the stylized facts highlighted in the literature, we find

[1]As can be attested by the excellent survey by Klenow and Malin (2009). See Bils and Klenow (2004), Dhyne, Alvarez, Bihan, Veronese, Dias, Hoffman, Jonker, Lunnenmann, Rumler, and Vilmunen (2005), Nakamura and Steinsson (2008), Bils, Klenow, and Malin (2009), Gagnon (2007), Gopinath and Rigobon (2008), Klenow and Kryvtsov (2008), Wulfsberg (2008).

[2]See Midrigan (2005) and Klenow and Kryvtsov (2008).

[3]In addition, some recent state-dependent models can also imply unimodal distributions. For example, a model with economies of scope in menu costs, such as Midrigan (2005)

[4]Cavallo (2010) found evidence of bi-modality using histograms in four Latin American countries, also included in this sample: Argentina, Brazil, Chile, and Colombia.

that the distribution of price changes is, in most cases, not unimodal at zero percent.

The data include individual-product prices in 36 supermarkets across 22 countries and 5 continents. They were collected by the *Billion Prices Project* (BPP) at MIT Sloan using a scraping software that records, on a daily basis, the price information for all goods sold by supermarkets with online shopping platforms.[5] These "scraped" prices were collected between October 2007 and February 2010, and on average, for each retailer we have 571 days of data, 20,000 products, 5 million price observations, and 80,000 price changes.

This "scraped" dataset has several advantages. First, we collected prices every day, as opposed to once a month (or two months) like most prices used in the CPIs. The daily data reduce the sampling biases that are associated with low frequency prices, as discussed later on. Second, we collected the full array of products sold by each retailer and therefore do not have forced item substitutions, nor do we rely on hedonics or other imputation procedures to compute prices – as occurs in some of the items underlying the CPI. There are also no adjustments for product discontinuations and technological improvements because these are rare in supermarket goods. All of these adjustments can potentially affect the size of price changes in other datasets. Finally, we collected posted prices as opposed to unit values. In most scanner datasets, prices are unit values computed as the ratio between total sales and total quantity sold at given frequencies (usually every week). Even daily unit values can experience small fluctuations due to different intensities in the use of loyalty cards, coupons, and quantity discounts which can introduce small price changes that are unrelated to the actual posted price change.[6]

The first part of our analysis uses the two best-known tests for unimodality available in the statistical literature: Hartigan's Dip and Silverman's Bandwidth tests. These tests are intuitive, easy to compute, and statistically powerful. We find that Hartigan's Dip rejects unimodality in 36 out of 37 supermarkets, while Silverman's test rejects the null of unimodality in 33 of those supermarkets. Although these results point to the rejection of unimodality, these tests have two limitations that complicate their interpretation. First, the main reason for these rejections is that these tests are too sensitive to even tiny bumps in the distribution, some of which may not be economically meaningful.[7] Our goal is to reject unimodality only when the distribution exhibits additional modes that are large enough to allow us to distinguish between different theories of price stickiness; as opposed to rejections

[5]For an introduction to Scraped Data, see Cavallo (2010)

[6]From the inflation calculation point of view, scanner data could have an advantage over the data we use. This would be the case if there are frequent changes in the use of discounts or other aspects of consumers demand, which are an important part of the welfare calculations of inflation that our data misses.

[7]In fact, when we apply Silverman's test to a null of bimodality – with an alternative hypothesis of more than two modes – it still rejects it for 14 supermarkets.

due to small jolts in the distribution. Second, both tests are not designed to measure modality around a specific value, like zero percent; hence, their rejections might be occurring around a point in the distribution that is far from zero, or at least far from the price change around which the economic analysis is focused.

To deal with these limitations we developed a new test called the *Proportional Mass (PM) test*. It was designed to find unimodality around a specific value, like zero percent, and to allow for small modes in the distribution as part of the null-hypothesis. The intuition of the test is the following: First, it computes the mass of price changes in absolute terms within certain bounds. These bounds form progressively larger intervals from the center, for example, at ±1% and ±5%. In a unimodal distribution, the mass in the smaller interval (close to the center) is larger than the proportional (per unit) mass in the larger interval. In a bimodal distribution (with two significantly large modes away from the center) the opposite is true. The PM score shows the relative importance of the mass close to the center. A higher PM score is evidence in favor of unimodality. This is a more conservative test because it requires the different modes to be of relatively similar importance in order to affect the relative masses. A distribution with one large mode and a series of smaller bumps will exhibit a similar proportional mass than a purely unimodal distribution. Intuitively, the test is not rejected for a multimodal distribution if the masses of the smaller modes can be rearranged to form a unimodal distribution with the same mass within that interval.[8]

Three results are worth highlighting. First, in two-thirds of the supermarkets we reject unimodality around zero percent. In particular, when we perform the PM test at the store (aggregate) level in the -5 to +5 percent interval, we reject unimodality at zero percent in 22 out of 36 supermarkets. This level of rejection is similar when we compute PM scores on a quarterly basis for each retailer. Second, heterogeneity at the category level can account for about half of the rejection of unimodality at the store level. Third, the data is consistent a model that combines elements of time and state-dependent pricing behaviors.

The first result is mostly at odds with the existing literature based on CPI and scanner datasets, so we explored possible explanations to reconcile the differences. We found two

[8] One possible source of concern in our test is the fact that *sales* can produce modes in the distribution that would lead us to reject unimodality. For example, imagine that the company does frequent temporary sales of 10 percent. In this case, the distribution of price changes will have symmetric modes around the negative and positive values of -10 and 10. In order to deal with this problem we focus our modality analysis in the -5 to +5 percent price change window, because reported and unreported *sales* smaller than 5 percent are rare in our sample. In addition, we expect adjustment costs that create any bimodality to have their largest impact within this range of values. As we show below, unimodality is more easily rejected if the full distribution is used. As an alternative, we can exclude sales using explicit indicators posted by the retailers online. However, only a fraction of retailers have sale indicators that we can collect, and even in those some cases, it is possible that not all sale events are explicitly identified.

reasons for the discrepancies with scanner data findings. First, scanner data tend to have "unit values" instead of actual posted prices. Stores report the total sales and total quantities per item, and unit values are computed as the ratio between these two totals. Unit values are affected by shifts in consumer purchasing practices. For example, consumers might decide to buy with or without coupons, or with or without loyalty cards. Therefore, the unit values can have small changes due to randomness in consumer demand and not because the posted price has actually shifted. Second, scanner data are usually reported on a weekly basis, so there is also an averaging that takes place through out the week. Although in our data we do not have prices with loyalty cards, we can simulate a weekly averaging or unit value. When we take our data and average the weekly prices, we fail to reject unimodality in 32 out of 37 supermarkets. A more challenging task is to reconcile our results with those underlying the CPI data. In the case of the US, our results are not necessarily at odds with the ones found in the CPI data, because two out of three supermarkets in our US sample are indeed unimodal. Still, there are important differences between our results and those found in other countries with monthly CPI data. Sampling frequency differences do not seem to offer a valid explanation: when we re-sampled our data to replicate the monthly sampling from statistical offices we find weaker results, but not weak enough to reduce the number of rejections of unimodality. Another possibility is the fact that statistical offices sometimes impute missing values with hedonic estimates, and average changes for similar goods in cases of forced substitution, discontinuations, and out-of-stock items.[9] Unfortunately, we do not have access to the CPI data to determine how common these practices are, and therefore we must leave this important question for future research.

The rejection of unimodality at the store level can be partially driven by category-level heterogeneity. Intuitively, it is possible that each category is unimodal by itself, but that each mode occurs at different average price changes, so that the aggregate distribution appears multimodal. To explore this possibility, we re-estimated the PM score for the narrowest categories of goods available in each supermarket. These categories are chosen by the supermarkets to display similar products on a single webpage, and we use the narrowest level of aggregation, such as "organic eggs" and "fat-free milk".[10] Moving to these categories substantially reduces the number of observations, making the PM test less powerful. Still, after running the test for each category, we find that on average across retailers we reject

[9] See BLS (2009). From the inflation measurement point of view, the computation of hedonics and the imputation of missing variables using statistical methods is the correct procedure. These practices, however, might lead to a pricing behavior that is not completely reflective of the posted prices by firms

[10] The categories do not exactly match the ten-digit Harmonized Tariff Schedules (HTS) from the US Census, but they are just as detailed. In fact, in developed countries, our supermarkets tend to organize their products in an average of 2000 categories, close to the 2500 categories for food and beverages at HTS (which also includes live animals, trees, etc).

unimodality in 30 percent of all categories. In supermarkets where bimodality was found at the store level, this level of rejection rises to 48 percent. Heterogeneity is indeed playing a role in the aggregate rejection of unimodality results, but it is not strong enough to make it disappear.[11]

Finally, we perform a simple simulation exercise with the model by Alvarez, Lippi, and Paciello (2010). We estimate the PM score from the simulated data at various levels of observation and adjustment costs, and find that the typical PM score in the data is consistent with a mixture of time and state dependent behaviors in the model. This exercise also illustrates how the PM test can be used to provide a unique statistic of the modality of empirical distributions that theoretical models can attempt to match.

The paper is organized as follows: Section 2 describes the data. Section 3 introduces three non-parametric statistical tests of unimodality. Section 4 presents the results of these tests, with evidence rejecting unimodality at zero percent. We also discuss some explanations for the difference in our results with the rest of the literature, and explore the effects of category-level heterogeneity. Section 5 simulates the Alvarez, Lippi, and Paciello (2010) model for different parameters and computes the PM test for unimodality in the simulated price change distributions. Section6 concludes.

2 Data: The Billion Prices Project

The data was collected by the *Billion Prices Project* (BPP) at MIT Sloan. We used a scraping software to record, on a daily basis, the price information for all goods sold by online supermarkets.

The scraping methodology for each retailer works in 3 steps: First, at a given time each day we download all public web-pages where product and price information are shown. These pages are individually retrieved using the same URL or web-address every day. Second, we analyze the underlying code and locate each piece of information that we want to collect. This is done by using custom characters in the code that identify the start and end of each variable, matching the format of that particular page and supermarket. For example, prices are usually shown with a dollar sign in front of them and two decimal digits at the end. This set of characters can be used by the scraping software to identify and record the price every

[11] At this moment we cannot match categories across supermarkets, but an interesting question is to determine which characteristics of the categories – if any – produces bimodality or unimodality. Unfortunately, even though we have the description of the categories, there are no classification standards across supermarkets, which makes this matching impossible. Categories in different supermarkets aggregate different products and there are often some overlaps.

day. Third, we store the scraped variables in a panel database, containing one record per product-day. Along with the price and product characteristics, retailers show an id for each product in the page's code (typically not visible when the page is displayed to the customer), which allows us to uniquely identify each product over time.[12]

The retailers included in this paper are detailed in Table 1. There are 36 supermarkets in 22 countries and 5 continents. Prices were collected on a daily basis between October 2007 and February 2010, with different starting dates for each supermarket. In all cases, we have at least one year of data, with a mean per retailer of 571 days, 20 thousand individual products, 5 million daily observations and 100 thousand price changes. When computing price changes, we consider only consecutive price observations, for which data is directly observed at days t and t-1.[13]

[Table 1 about here]

3 Tests for Unimodality

The standard analysis of unimodality in the micro-price literature relies on histograms and cumulative frequency plots.[14] This is adequate to examine the shape of few distributions, but not practical to extend the analysis to a large number of retailers and countries.

We formally test for unimodality using three non-parametric statistical tests: Hartigan's Dip Test, Silverman's Bandwidth Test, and a test we develop in this paper called the *Proportional Mass Test*.

Hartigan's and Silverman's tests are common in the statistical literature but have rarely been used in economic applications before, so we describe them briefly below.[15] They are

[12]For more on the scraping methodology, see Cavallo (2010) and www.billionpricesproject.org

[13]The collection of high-frequency information for every single product sold in each supermarket greatly expands the number of data points available, but at the same time, produces frequent gaps in individual price series. These gaps occur, for example, when the scraping software fails or individual items are temporarily out of stock. Scraping failures are typically resolved in a few days by the BPP scraping team, but seasonal products can create missing values that last several months. The standard treatment in the literature is to fill missing prices with the last recorded price available for each product, but if some price changes are not directly observed, the distribution of the size of changes can be affected. For example, in cases of high inflation, price changes could appear larger, because several price adjustments would be accumulated over time. By contrast, in a context with low inflation but frequent temporary shocks, two unobserved price changes of opposite magnitudes could appear to be one smaller change in the data. We take a conservative approach in this paper and consider only consecutive price observations as the basis for computing price changes.

[14]See Kashyap (1995), Klenow and Kryvtsov (2008), Kackmeister (2007), Midrigan (2005) and Cavallo (2010)

[15]See Hartigan and Hartigan (1985) and Silverman (1981). A recent paper that uses both tests in an

both intuitively appealing and simple to compute. Unfortunately, they are also very sensitive to small modes in the distribution that may not be economically relevant but still lead to frequent rejections of unimodality. In addition, these tests can only take into account the full distribution. So, for example, if are two modes, one centered at 0% and the other at 5%, both tests would reject unimodality in general, but tell us nothing about the modality around 0%, which is key to evaluate the prediction of time-dependent and state-dependent models.

By contrast, our *Proportional Mass Test* (PM) has two main advantages in the context of a micro-price setting paper: first, it ignores tiny modes that may not be economically meaningful, and second, it tests for modality in an interval around a specific value, such as zero percent.

3.1 Hartigan's Dip Test

The dip test, described in Hartigan and Hartigan (1985), relies on the fact that the cumulative distribution function of a density function f with a single mode at m_f is convex on the interval $(-\infty, m_f)$ and concave on the interval (m_f, ∞). The intuition of this property is simple: at the right hand side of the mode, the density is non increasing – meaning that its derivative is non-positive. The opposite occurs at the left of the mode.

Using this property, the dip statistic is calculated to measure the departure of an empirical distribution from the best fitting unimodal distribution. When the empirical distribution has a single mode, the dip statistic is zero. If the empirical distribution has multiple modes, with a cumulative distribution that has several regions of convexity and concavity, then it will be "stretched" until it takes the shape of an unimodal distribution. The larger the stretch needed, the larger the departure from unimodality and therefore the larger the *dip* becomes.

In Hartigan's test, positive dip values provide evidence to reject the null hypothesis of unimodality. To determine the statistical significance of a positive dip, the test sets the null hypothesis equal to the uniform distribution, for which, asymptotically, the dip value is stochastically largest among all unimodal distributions.[16] This increases the power of the test, making it more likely to reject the null hypothesis of unimodality.

economics setting is Henderson, Parmeter, and Russell (2008). *Parametric* tests of modality are more common in economics. Examples of these tests include Paapaa and van Dijk (1998) and Anderson (2004), with methods that mix normal distributions with mass overlaps. Unfortunately, these parametric tests require the *ex-ante* assumption of a number of clusters or groups, and are used to reject the null hypothesis of normality, not general forms of unimodality.

[16]Hartigan and Hartigan (1985) argue that this is not always the case with small samples, so we use a calibration of the dip test proposed by Cheng and Hall (1998) to account for such cases

3.2 Silverman's Bandwidth Test

Silverman's Bandwidth or "Bump" test uses kernel smoothing functions to evaluate modality. Given a sample $X = (x_1, x_2, ..., x_n)$, a non-parametric kernel estimate of the unknown density function f is given by

$$\hat{f}(x,h) = (nh)^{-1} \sum_{i=1}^{n} K\left(\frac{x_i - x}{h}\right) \tag{1}$$

where h is the smoothing parameter (or "bandwidth") and K is the Gaussian kernel function. Silverman (1981) showed that the larger smoothing h, the fewer the number of modes in $\hat{f}(x,h)$. Therefore, for the null hypothesis of unimodality, he proposed the test statistic

$$\hat{h}^1_{crit} = \inf\left\{h : \hat{f}(x,h) \text{ has 1 mode}\right\} \tag{2}$$

This is the minimum smoothing required for the smoothed kernel density to have one mode. Large values of $\hat{h}^1_{crit}$ are evidence against the null hypothesis, because larger degrees of smoothing are needed to eliminate additional modes in the density estimate.

The statistical significance of $\hat{h}^1_{crit}$ is evaluated using a smoothed bootstrap method.[17] For each bootstrapped sample, we compute the minimum bandwidth $\hat{h}^{1*}_{crit}$ required to have one mode and estimate the probability $\hat{P}$, given by

$$\hat{P} = P\left(\hat{h}^{1*}_{crit} \geq \hat{h}^1_{crit}\right) \tag{3}$$

$\hat{P}$ gives us a way to know the relative level of $\hat{h}^1_{crit}$. If it is relatively high compared to the results from the bootstrapped samples, then $\hat{P}$ will be small and there is stronger evidence against the null hypothesis.[18]

This method can be used to test for any number of modes, and is usually carried out in sequence, starting with one mode and continuing until the test fails to reject the null hypothesis of m modes. This is a major advantage of Silverman's approach, because it allows us to test explicitly for bi-modality in the size of price changes. In addition, this test

[17]The bootstraps are drawn from an smoothed conditional function re-scaled to have a variance equal to the sample variance. See Henderson, Parmeter, and Russell (2008) for details.

[18]Because the number of modes is non-increasing with h, $\hat{P}$ is equivalent to the share of bootstraps that have more than one mode when evaluated with bandwidth $\hat{h}^1_{crit}$. We use this approach to estimate $\hat{P}$, also called the *achieved significance level* in the bootstrap literature, because it is easier to compute.

is intuitively appealing and easy to compute. Unfortunately, it also has important weaknesses derived from the use of a single bandwidth in the kernel smoothing estimates: it is easily affected by outliers in the tails of the distribution and it is sensitive to tiny bumps which lead to frequent rejections of the null hypothesis, especially in large samples.

3.3 Proportional Mass Test

We now propose a new "Proportional Mass Test" that compares the relative mass of the distribution between bounds to determine the degree of unimodality around a centered value.

The test relies on the fact that unimodal distributions have a high proportion of their mass close to the mode. If we take an interval around the mode and make it progressively larger, the total mass increases by *smaller* increments each time. By contrast, in a distribution that is bimodal around a specific value, the mass increases by *larger* increments each time. Therefore, the relative size of these additional increments of mass can be used to determine the degree of unimodality in the distribution around a specific value.

To illustrate this, consider the case where the distribution is unimodal centered at zero percent, as shown in Figure 1(a). The mass between -1% and 1% should be larger than the mass between -5 and 5 per unit, that is,

$$P\left(|\Delta p| \leq 1\right) \geq P\left(|\Delta p| \leq 5\right)/5 \tag{4}$$

The *proportional mass* between $i = 1$ and $j = 5$ is thus given by

$$PM_{1,5}^{0} = \ln \frac{P\left(|\Delta p| \leq 1\right)}{P\left(|\Delta p| \leq 5\right)/5} \tag{5}$$

This ratio is positive when the distribution is unimodal around zero.[19] By contrast, when the distribution is strictly bimodal around zero, $PM_{1,5}^{0}$ is negative, as shown in Figure 1(b). Finally, if the distribution is bimodal but the modes are not large, as seen in Figures 1(c) and 1(d), then the PM will remain positive. This ensures that minor bumps in the distribution will not cause a rejection of unimodality.

The ratio can be generalized to incorporate information from different intervals and compute the *Proportional Mass Score* around zero, given by

[19] If the distribution is uniform, $PM_{1,5}^{0} = 0$ when the domain of the distribution is wider than 5, otherwise $PM_{1,5}^{0}$ is positive

$$PM^0 = \frac{1}{|Z|} \sum_{ij \in Z} PM_{ij} \tag{6}$$

where Z is the set of all combination ij such that $i < j$.

The same logic applies when we want to test the degree of unimodality around a mode m, with PM^m given by

$$PM^m = \frac{1}{|Z|} \sum_{ij \in Z} \ln \frac{P\left(|\Delta p - m| \leq i\right)}{P\left(|\Delta p - m| \leq j\right)/(j/i)} \tag{7}$$

In our computations, we consider the intervals $i, j \in \{1, 2.5, 5\}$, but we also test the robustness of our results to changes in these intervals.[20]

The null hypothesis of the test is that PM^m (the "PM score") is positive, corresponding to a unimodal distribution. The statistical significance is evaluated using bootstrapped samples from the data and calculating the share with positive PM scores. The lower the share of bootstraps with positive PM scores, the stronger the rejection of unimodality.

4 Results

4.1 Rejection of Unimodality at 0%

We first ran Hartigan's Dip test in all supermarkets. The first two columns in Table 2 show the dip statistics and p-values for the null hypothesis of unimodality. The dip statistics are consistent with a simple graphical analysis of the histograms in Figures 2 to 4. For example, the lowest dips belong to AUSTRALIA-4, NETHERLANDS-1, UK-1, UK-2, UK-3, and COLOMBIA-1. These are cases that either uniformly distributed or have a large dominating mode.

As a statistical test, Hartigan's method is just too powerful for our purposes. At the 1% significance level, unimodality is rejected in 36 out of 37 supermarkets. The test rejects the null hypothesis even for distributions with only minor departures from unimodality.

[20] A key parameter in the design of this test is the optimal bandwidth. It is likely to depend on the type of adjustment cost, and the reason behind the price bumps in the distribution. In this paper we focus on the -5% to 5% price change range to avoid modes generated by sale events, and also because adjustment costs that could create bimodality are likely to have their greatest impact in this range of price changes. The aspects of optimal design for the PM test are left for future research.

[Table 2 about here]

Next, we considered Silverman's bandwidth test. The results are shown in columns 3 to 5 of Table 2. The critical bandwidth values, which measure the degree of "smoothing" needed to obtain a single-mode kernel estimate, are also consistent with a simple graphical analysis. Some of the lowest critical values are, once again, in AUSTRALIA-4, UK-1, UK-2, UK-3 and COLOMBIA-1.

Although slightly more conservative, Silverman's test still rejects the null of unimodality in 33 out of 37 supermarkets. The rejection level is high even when we consider the null hypothesis of 2 or less modes. In fact, in 22 supermarkets we find evidence supporting *more* than 2 modes. The test also appears to be too sensitive to tiny bumps in the distribution. This is especially true in those retailers with the largest number of observations, such as URUGUAY-1, CHINA-2, CHILE-1, RUSSIA-1, IRELAND-1, US-1 and NEWZEALAND-1, where we reject both unimodality and bimodality around zero.

As we mentioned before, we consider the excessive sensitivity of both of these tests to small bumps in the distribution as a major weakness. We are looking for modes that are sufficiently large and can provide insights into the importance of menu costs and other pricing behaviors. So we move to analyze the results from the PM test, which is significantly more conservative.

The results for the PM test centered at 0% are presented in Table 3. Column 3 shows the PM score point estimate, columns 4 and 5 show the mean and the standard deviation in 500 bootstrapped samples, and column 6 shows the share of bootstrapped estimates that have a positive PM score (unimodality).

[Table 3 about here]

As expected, the PM test is far more conservative. We now fail to reject unimodality in 14 supermarkets, or 40% of the total. The reason is that this test does a better job at ignoring small bumps in the distribution, because it spreads their mass into relatively wide intervals that are used to calculate the proportional mass ratios.

Comparing the PM results to Silverman's test, and the graphical distribution in Figures 2 to 2, we can see why the PM score is a better test for your purposes. For example, the largest PM scores belong to the distributions of URUGUAY-1, UK-2, UK-1, and US-1, all of which are identified now as unimodal. By contrast, Silverman's test results suggested they were all bimodal or multimodal. In the case of URUGUAY-1, where nearly all the mass lies within -1% and 1%, Silverman's test rejected unimodality because there is (by construction), no mass at exactly zero percent, and this has a great impact on the smoothed kernel for this tiny range of values. At the same time, the cases of UK-2, UK-1, and US-1 clearly do have a single big mode, but lots of tiny modes that also cause Silverman's test rejection of unimodality.

Still, even though we have been stacking the odds to find some unimodality, the PM test continues to rejects the null hypothesis in 22 supermarkets, about two-thirds of the total. The evidence against unimodality at zero percent is simply very strong in our data. This finding is robust over time, as seen in In Table 4. When we compute the PM score at quarterly intervals, we still obtain negative scores (bi-modality) in most retailers over time.

[Table 4 about here]

Overall, the three statistical tests strongly reject the hypothesis of unimodality around zero percent. We have shown results within the interval of -5% to 5%, but these findings are robust to extensions with distributions at the +/- 10% and +/- 50% intervals. In fact, the wider the range of the distribution, the lower the evidence of unimodality around zero percent.[21]

4.2 Reconciling differences with the Literature

Our main finding, the lack of unimodality of price changes around zero percent, is at odds with the existing literature that uses Scanner and CPI data. In this section, we consider possible explanations for these differences by replicating some of the sampling methodologies in these two types of data.

[21] While there is less unimodality around zero percent, there can still be large modes away from zero. This In the Appendix, we explore this possibility by running the PM test centered on the highest "mode" in the distribution. Positive PM scores indicate the presence of modes that are large enough to dominate the mass of price changes within a +/-5% interval of that value, possibly reflecting the outcome of an inflationary or deflationary macroeconomic context. In Table A2, we find that the highest mode is negative in 13 supermarkets and positive in 21 supermarkets, and 34 out of 36 supermarkets have a *positive* PM score. Computing PM scores away from zero can also be used to explore the changes in modality with different levels of inflation, which has important implications for some theoretical models. For example, standard state-dependent models predict that an economy that moves gradually from a peak of inflation to a peak of deflation will have a distribution that looks initially unimodal with a positive mode, then bimodal at zero, and finally unimodal with a negative mode. Table A3 shows the share of bootstraps with a non-unimodal PM score away from zero in each quarter. We find that the distributions were less unimodal in late 2008 and early 2009. The last row shows that the ratio of retailers with some evidence of bimodality starts to rise in the fourth quarter of 2008 and peaks in the second quarter of 2009. This is a time when many of these countries were in the middle of a recession. Although the shift in modality is not as stylized as standard models predict, these results suggest that modality and inflation could be closely related.

4.2.1 Differences with Scanner Data

Scanner datasets have two important differences with our data. First, prices are constructed as "unit values", with the ratio of total sales over total quantity sold for each product. Because consumers can sometimes purchase products with or without coupons, with or without loyalty cards, or even at different prices within the same day, this unit value will change in small percentages with the randomness in consumer demand. Second, scanner data are reported on a weekly basis, so there is also an averaging that takes place along the week. The effect of this averaging is discussed by Campbell and Eden (2005). Their focus was not on the size of changes, but they described some complications caused by weekly averages using a simple example. Consider a three week period with a single price change on the middle of the second week. If average weekly prices are used, each week would have a different price and two –smaller– price changes would be observed.

Although we do not have information on the use of loyalty cards and coupons, we can replicate the weekly averaging in our data and see how it affects our results. We do so by first computing the weekly average price per individual product, and then re-calculating price changes only when consecutive weekly prices are available.

Our results in Table 5 show that the evidence of unimodality increases dramatically with weekly averaged prices. This table compares the effect of weekly averaging on the three measures of modality embedded in our tests: the dip statistic, the critical bandwidth and the PM score (centered at 0%). A drop in Hartigan's dip means that, on average, the distribution is now closer to being unimodal. A drop in Silverman's critical bandwidth means that less smoothing is needed to obtain an unimodal kernel estimate. An increase in PM scores means that the distribution becomes unimodal around zero. In all three cases, the evidence for unimodality increases dramatically with weekly prices. Furthermore, the PM test centered at zero also fails to reject unimodality in 32 out of 37 supermarkets.

[Table 5 about here]

4.2.2 Differences with CPI Data

Reconciling our results with CPI studies is harder because the differences in the data go far beyond simple sampling methodologies. Nevertheless, the monthly sampling of prices could lead to artificially small price changes when there frequent temporary shocks lasting less than a month. For example, if a price were to fall from \$10 to \$9, and then move back to \$10.1 within a few days, monthly sampling would detect a +1% price change instead of two changes of -10% and +12%. Cavallo (2010) showed that these type of temporary changes can occur frequently in supermarket data, and it can be particularly relevant in low-inflation settings like the US, where most of the literature's CPI findings come from.[22]

To approximate the CPI sampling methods, we randomly picked one day of the month for each individual product and recorded the price. If we chose a day where no price information is available, the price is missing for that month. Next, we re-calculated price changes only when consecutive monthly price observations were available.

In contrast to weekly averages, monthly sampling of the data has no effect on the degree of unimodality. The average dip statistic, critical bandwidth and PM score in Table 5 are similar with daily and monthly data (even though the number of observations drops significantly once monthly data is used).

An alternative explanation for our differences with the CPI literature it related to individual price corrections in the US CPI series. The BLS makes several adjustments in individual prices that can potentially affect the distribution of the size of changes. First, changes in a price spell can be cause by *forced item substitutions* that occur when an item is no longer available. In these cases, the BLS estimates a price change using hedonic quality adjustments or the average price change for that category of products. Second, even when no product substitutions occur, the BLS sometimes imputes prices that are considered temporarily missing. Seasonal products –including Fresh Food– are the typical case when this happens. Third, individual prices can also be adjusted for coupons, rebates, loyalty cards, bonus merchandize, and quantity discounts, depending on the share of sales volume that had these discounts during the collection period. Fourth, some food items that are sold on a unit basis –like apples– are sometimes weighted in pairs to calculate an average-weight price. These and other price adjustments are described in the BLS Handbook of Methods.[23] Unfortunately, we do not know how frequent these changes are in practice, or whether they can explain most of the small price adjustments previously found by the literature. Without access to the US CPI data, we must leave this important questions for future research.

4.3 Product and Category Heterogeneity

The existence of multiple modes documented so far at the store level could be the result of heterogeneity across categories. It is possible that each individual distribution is unimodal -at different values- and the aggregate exhibits more than one mode. In this section we

[22]In a setting with high inflation, monthly sampling can have the opposite effect, accumulating several small price changes that occur within a month.

[23]See , Chapter 17, pages 30 to 33.

explore the role that heterogeneity in the PM test results.

In the process of collecting the data, we record the category in which the supermarket placed each item. These categories are constructed by each retailer to facilitate the finding of individual items in the website. For example, a typical supermarket would show a navigation menu with "Food", "Diary", "Milk", and finally "Low-Fat Milk". We use the narrowest level of aggregation, such as "Low-Fat Milk" in this example.[24] There are several advantages in this categorization. First, products are added and reclassified automatically by each retailer. Second, items within each of these categories are close substitutes, because classification depends more on consumer's practices than on a preconceived notions of which items we consider to be closely related. For example, when iPad was launched in 2010, it was unclear whether it would compete mostly with laptops, netbooks, or e-book readers. As other tablets appeared in the market, a new category was created by retailers. This dynamic categorization ensures that we can compare pricing behaviors between close competitors. The drawback is that international comparisons between categories are not possible, because stores in different countries tend to arrange products in different ways.

In each category, we first obtain 500 bootstraps to estimate the distribution of the PM test – the same procedure we followed at the supermarket level before. We then compute the proportion of bootstrapped PM scores that are negative – indicating a rejection of unimodality. Finally, for each supermarket, we calculate the proportion of categories where unimodality is rejected in 95 percent of the bootstrap replications, in 90 percent, 85 percent, and so on. These results are shown in Table 6. The last two columns indicate the total categories available in each store, and the total number of categories in which there are at least 64 price changes, enough to run the test. These represent only about 20% of the total number of categories.

[Table 6 about here]

The first column shows the proportion of categories in which 95 percent (or more) of the replications in the bootstrap rejected unimodality. As can be seen, there are some supermarkets in which all the categories rejected unimodality, while there are others in which not a single one of the categories rejected unimodality. However, in most supermarkets, there are approximately 40 percent of categories where we can reject unimodality. Because not all supermarkets are of the same size, the *weighted* average proportion of rejections (weighted by the number of categories where the test can be run) is 27.1 percent.

[24]See Cavallo (2010) for a more thorough description of these categories in four of the countries in our sample.

Notice that the rejection at the category level is much smaller than the rejections we reported at the supermarket level in previous sections (over 60%). Besides heterogeneity, it is possible that, because we have less observations, the PM test looses too much statistical power. To evaluate this possibility we can move along the columns of Table 6, reducing our confidence interval. For example, in the third column we present the proportion of categories in each retailer in which at least 85 percent of the simulations rejected the null hypothesis of unimodality. In this case, the average unweighed proportion is 43.8 percent, and the weighted one is 31.3 percent. Although larger than before, the increase in rejections is not overwhelming. In fact, we need to reduce the confidence interval to 75 percent (i.e. a significance level of 25 percent) to get similar unweighed rejections as in the previous section.[25]

Overall, we believe that part of the rejections that we observed at the supermarket level are due to category heterogeneity. There is still a significant proportion of rejections of unimodality in the data, at least 30 percent on average.

Finally, could the bimodality in these categories be caused, in turn, by product-level heterogeneity? This is unlikely in theory, precisely because products classified into these narrow categories are supposed to be close substitute. So, for example, it is reasonable to expect products in the "organic eggs" category to be quite homogeneous in their pricing behavior. Still, while we cannot run the PM test at the individual-product level due to lack of data (very few items have 64 price changes), we did not want to leave the possibility of product heterogeneity unexplored, so we explored a simple question: does the dispersion of inflation rates at the item level increase the likelihood of rejecting unimodality. If we assume that the mode is closely related to the average inflation, then this exercise is meaningful: categories where the average inflation is very different across products are categories where rejection of unimodality should also be more likely. We understand we are making a strong assumption: that the only heterogeneity that matters is the average inflation. However, computing the average daily inflation rate in one item is straightforward, because it does not depend on the number of observations but on the span of the price collection.

[Table 7 about here]

Table 7 presents the result of a series of simple regressions. Within each supermarket,

[25] We could have a selection problem if there is a relationship between the cause for stickiness and the frequency of price adjustment. For example, suppose that products that are more time-dependent (and therefore more unimodal) tend to have more frequent price adjustments. If that were the case, then by requiring categories to have a minimum number of observations we are biasing the sample towards unimodality. However, when we run the exact same exercise on categories with at least 32 price changes, our results barely change: the unweighed proportion of rejections is 38.9 percent, while the weighted is 27.8 percent.

we put on the left hand side the proportion of simulations that reject unimodality in a category, and on the right hand side the inflation dispersion within that category. There is one observation per category. We estimate a simple correlation, estimated as an OLS regression where a constant is included. For each supermarket in which we were able to estimate this regression we present the coefficient (column 2), its standard deviation (column 3), and the t-stat (column 4).

The purpose of this exercise is to determine if the rejection of unimodality in a category (higher proportion of simulations rejecting unimodality) is correlated with product heterogeneity – in this case measured as inflation dispersion. Hence, a positive coefficient is evidence in support of heterogeneity causing the rejections. Only 23 supermarkets have enough observations for us to perform the estimation. From them, 12 have statistically significant positive coefficients, 11 are not significant; and only 3 are negative (but not significant). So, in about half of the supermarkets the rejection can be partially explained by product heterogeneity.

Overall, this is consistent with the message of heterogeneity we highlighted when studying the PM scores at the category level. We loose observations and statistical power, but in those cases were the test can still be run, we find that heterogeneity is playing an significant role. Future research – with several more years of individual data – should further address the question of product-level heterogeneity.

5 Simulation of a model with both state and time-dependent pricing

The results so far show that there is evidence of both bimodality and unimodality in the size of price changes around zero percent. In this section we simulate a model that exhibits menu cost and observation/information costs to evaluate the strength and properties of the PM test, and to be able to make an assessment of the relative importance of both time and state-dependant pricing behaviors. We use the recent model by Alvarez, Lippi, and Paciello (2010). They assume a firm solving a pricing problem, with quadratic cost function, exhibiting fixed cost to change prices, and a fixed cost to observe previous realizations. This is a stylized model that has relatively simple solutions. The fixed cost for changing prices is the standard menu cost, while the observation cost alone can generate an optimal strategy that resembles the time-dependent rule in the Calvo (1983) model.[26] The advantage of Alvarez, Lippi, and Paciello (2010) is that encompasses both types of costs, and can also account for

[26]See Mankiw and Reis (2002) and Mankiw and Reis (2007)

different levels of inflation.

We simulate the model under zero inflation and use a range of menu costs from 0.4% to 1% and an observation cost from 1% to 6%, consistent with the range selected by Alvarez, Lippi, and Paciello (2010).[27] For each of the simulations, we compute the distribution of price changes and estimate the PM score. Figure 5 computes the PM test when the 0 to 5 percent window is used.[28] There are two panels in this figure. The top panel shows the surface for several choices of menu cost and observation cost. The bottom panel is an iso-PM score figure – the combination of menu and observation costs that produce the same PM score within the range we studied.

The PM score is, as expected, affected by the size of each cost. First, an increase in the menu cost reduces the PM score unambiguously. Second, an increase in the observation cost increases the PM score also unambiguously. Both of these implications should be expected. When the menu costs are increasing the range of inaction for the firm increases, reducing its mass around zero, and making the distribution more bimodal. On the other hand, increasing the observation costs imply a behavior closer to the standard Calvo model, and therefore the distribution of price changes becomes more unimodal, with a higher PM score.

[Figure 5 about here]

The PM score could be used in this model to get a sense of the magnitude and relative importance of the observation and adjustment costs. For example, Figure 5 tells us that if the PM score lies between -0.5 and -0.6, then the menu cost would have to be relatively small – from 0.6% to 1% – while the observation costs about 3 to 5 times larger –1.5% to 5%–. This is consistent with the estimates in Alvarez, Lippi, and Paciello (2010)'s own calibrations. The menu costs, in particular, are also close to the estimate of 0.7% of revenue obtained by Levy, Bergen, Dutta, and Venable (1997), who looked at direct evidence of menu costs in a large US supermarket chain in the 90s.

6 Conclusions

The distribution of the size of price changes is an important implication of the different theories of price stickiness. One of the key characteristics of this shape is the number of modes

[27] The model is estimated with the additional parameters: $B = 20$ (cost function parameter) as in Alvarez, Lippi, and Paciello (2010), a standard deviation in the target price $\sigma = 0.008$, and a daily discount factor $\rho = 0.0007$.

[28] See the Appendix Figure A1 for a wider window of 0 to 20 percent in PM Scores

around zero percent. We formally tested for this modality in a large set of supermarkets, spanning 22 countries and 5 continents, using the two best-known tests in the statistical literature –Hartigan's Dip and Silverman's Bandwidth– and a test designed in this paper –the Proportional Mass test–. When the traditional tests are used, the unimodality around zero is rejected in about 90 percent of the establishments. When we use the Proportional Mass test, which is much more conservative than the first two, we still reject unimodality in two thirds of the supermarkets. Heterogeneity at the product-category level can account for part of the rejections of unimodality at the retailer level.

While the rejection of unimodality implies an important role for adjustment or "menu" costs, there is no conclusive evidence in favor of any standard theory of price stickiness. The distributions we observe with the data appear similar, in fact, with those predicted by models that combine elements of both time and state-dependent pricing. We have shown how the PM score can be used to explore issues like these in a particular model, from Alvarez, Lippi, and Paciello (2010), but the relative importance of adjustment and observation costs is still an open empirical question whose answer is specific to the model being used.

We believe the PM test developed here will help this literature by providing a simple mechanism to evaluate modality in a large set of countries and retailers. Further research is needed to understand how different theories can explain cross-country and cross-retailer differences in PM scores and modality results, how they change through time, and across product categories. In particular, the link between inflation rates, the distribution's symmetry and the location of its modes seems a promising area for future work. In addition, the differences with CPI datasets should be further studied in countries where both micro-CPI and scraped data become available. Finally, while we do not have enough data today to compute PM scores at the individual product level, future research should study how heterogeneity within categories can impact our main results.

References

ALVAREZ, F. E., F. LIPPI, AND L. PACIELLO (2010): "Optimal price setting with observation and menu costs," *National Bureau of Economic Research Working Paper Series*, No. 15852.

ANDERSON, G. (2004): "Toward an empirical analysis of polarization," *Journal of Econometrics*, 122(1), 1–26.

BAJARI, P., AND A. HORTACSU (2003): "Economic Insights from Internet Auctions: A Survey," *National Bureau of Economic Research Working Paper Series*, No. 10076, published as Bajari, Patrick and Ali Hortacsu. "Economic Insights from Internet Auctions." Journal of Economic Literature XLII (2004): 457-486.

BILS, M., AND P. J. KLENOW (2004): "Some Evidence on the Importance of Sticky Prices," *Journal of Political Economy*, 112, 947–985.

BILS, M., P. J. KLENOW, AND B. A. MALIN (2009): "Reset Price Inflation and the Impact of Monetary Policy Shocks," *NBER Working Papers*, 14787.

BLS (2009): "The Consumer Price Index," in *BLS Handbook of Methods*, vol. Chapter 17.

BONOMO, M., C. CARVALHO, AND R. GARCIA (2010): "State-Dependent Pricing Under Infrequent Information: A Unified Framework," *Working Paper*.

CALVO, G. (1983): "Staggered Prices in a Utility Maximizing Framework," *Journal of Monetary Economics*, 12, 383–398.

CAMPBELL, J. R., AND B. EDEN (2005): "Rigid prices: evidence from US scanner data," *Federal Reserve Bank of Chicago*, Published: Federal Reserve Bank of Chicago.

CARLTON, D. W. (1986): "The Rigidity of Prices," *American Economic Review*, 76(4), 637–58.

CAVALLO, A. (2010): "Scraped Data and Sticky Prices," *Working Paper*.

CECCHETTI, S. G. (1986): "The Frequency of Price Adjustment: A Study of Newsstand Prices of Magazines," *Journal of Econometrics*, 31, 255–274.

CHENG, M., AND P. HALL (1998): "Calibrating the Excess Mass and Dip Tests of Modality," *Journal of the Royal Statistical Society. Series B (Statistical Methodology)*, 60(3), 579–589, ArticleType: primary_article / Full publication date: 1998 / Copyright 1998 Royal Statistical Society.

DHYNE, E., L. ALVAREZ, H. L. BIHAN, G. VERONESE, D. DIAS, J. HOFFMAN, N. JONKER, P. LUNNENMANN, F. RUMLER, AND J. VILMUNEN (2005): "Price Setting in the Euro Area: Some Stylized Facts from Individual Price Data," *Working Paper Series, European Central Bank*, No. 524.

Eichenbaum, M., N. Jaimovich, and S. Rebelo (2008): "Reference Prices and Nominal Rigidities," *National Bureau of Economic Research Working Paper Series*, No. 13829.

Fisher, J. D. K. T. C. G. (2000): "Synchronization of Price Changes by Multiproduct Firms: Evidence from Canadian Newspaper Prices," *Economics Letters*, 68.

Gagnon, E. (2007): "Price Setting during Low and High Inflation: Evidence from Mexico," *International Finance Discussion Papers, Board of Governors of the Federal Reserve System*, 896.

Golosov, M., and R. Lucas (2007): "Menu Costs and Phillips Curves," *Journal of Political Economy*, 115(2).

Gopinath, G., and R. Rigobon (2008): "Sticky Borders," *Quarterly Journal of Economics*, 23(2).

Hartigan, J. A., and P. M. Hartigan (1985): "The Dip Test of Unimodality," *The Annals of Statistics*, 13(1), 70–84.

Henderson, D. J., C. F. Parmeter, and R. R. Russell (2008): "Modes, weighted modes, and calibrated modes: evidence of clustering using modality tests," *Journal of Applied Econometrics*, 23(5), 607–638.

Kackmeister, A. (2007): "Yesterday's Bad Times Are Today's Good Old Times: Retail Price Changes Are More Frequent Today Than in the 1890s," *Journal of Money, Credit and Banking*, 39(8), 1987–2020.

Kashyap, A. K. (1995): "Sticky Prices: New Evidence from Retail Catalogues.," *Quarterly Journal of Economics*, 110, 245–274.

Kehoe, P., and V. Midrigan (2007): "Sales and the Real Effects of Monetary Policy," *Federal Reserve Bank of Minneaplis Working Paper*, No. 413.

Klenow, P. J., and O. Kryvtsov (2008): "State-Dependent or Time-Dependent Pricing: Does it Matter for Recent US Inflation?," *The Quarterly Journal of Economics*, 73(3), 863–903.

Klenow, P. J., and B. A. Malin (2009): "Microeconomic Evidence on Price-Setting," *Handbook of Monetary Economics*, forthcoming.

Lach, S., and D. Tsiddon (2007): "Small price changes and menu costs," *Managerial and Decision Economics*, 28(7), 649–656.

Levy, D., M. Bergen, S. Dutta, and R. Venable (1997): "The Magnitude of Menu Costs: Direct Evidence from Large U.S. Supermarket Chains," *The Quarterly Journal of Economics*, 112(3), 791–825.

Mackowiak, B., and F. Smets (2008): "On Implications of Micro Price Data for Macro Models," *Conference Series, Federal Reserve Bank of Boston.*

Mankiw, N. G., and R. Reis (2002): "Sticky Information Versus Sticky Prices: A Proposal To Replace The New Keynesian Phillips Curve," *The Quarterly Journal of Economics*, 117(4), 1295–1328.

——— (2007): "Sticky Information in General Equilibrium," *Journal of the European Economic Association*, 5(2-3), 603–613.

Midrigan, V. (2005): "Menu Costs, Multi-Product Firms and Aggregate Fluctuations," *Working Paper*.

Muller, D. W., and G. Sawitzki (1991): "Excess Mass Estimates and Tests for Multimodality," *Journal of the American Statistical Association*, 86(415), 738–746, ArticleType: primary_article / Full publication date: Sep., 1991 / Copyright 1991 American Statistical Association.

Nakamura, E., and J. Steinsson (2008): "Five Facts About Prices: A Reevaluation of Menu Cost Models," *Quarterly Journal of Economics*, 123(4), 1415–1464.

Neiman, B. (2008): "Multinationals, Intrafirm Trades and International Macro Dynamics," *Job Market Paper - Harvard University*.

Paapaa, R., and H. K. van Dijk (1998): "Distribution and mobility of wealth of nations," *European Economic Review*, 42(7), 1269–1293.

Silverman, B. W. (1981): "Using Kernel Density Estimates to Investigate Multimodality," *Journal of the Royal Statistical Society. Series B (Methodological)*, 43(1), 97–99, ArticleType: primary_article / Full publication date: 1981 / Copyright 1981 Royal Statistical Society.

Woodford, M. (2009): "Information-constrained State-dependent Pricing," *Journal of Monetary Economics*, 56, S100 – S124.

Wulfsberg, F. (2008): "Price Adjustments and Inflation - Evidence from Consumer Price Data in Norway 1975-2004," *Norges Bank Working Paper*.

Tables

Table 1: Supermarket Data

Database	Country	Started	Days	Obs.	Products	# Pr P/day	Pr. Ch.	(cc)	Sales
ARGENTINA-1	Argentina	10/7/2007	876	13117K	26K	12K	155K	1.2%	YES
ARGENTINA-2	Argentina	23/7/2007	861	5294K	11K	6K	103K	2.0%	YES
AUSTRALIA-1	Australia	8/4/2008	574	232K	3K	1K	147K	63.4%	NO
AUSTRALIA-2	Australia	8/7/2008	571	202K	1K	0K	2K	1.0%	NO
AUSTRALIA-3	Australia	8/4/2009	209	3292K	7K	6K	2K	0.1%	NO
AUSTRALIA-4	Australia	5/3/2008	667	1967K	18K	4K	46K	2.3%	YES
BRAZIL-1	Brasil	10/10/2007	873	10780K	22K	11K	260K	2.4%	YES
CHILE-1	Chile	10/24/2007	859	12102K	35K	12K	120K	1.0%	NO
CHINA-1	China	12/5/2008	451	1101K	7K	3K	6K	0.5%	NO
CHINA-2	China	3/19/2008	712	6644K	46K	10K	22K	0.3%	NO
COLOMBIA-1	Colombia	11/13/2007	839	4186K	9K	5K	77K	1.8%	YES
ECUADOR-1	Ecuador	3/19/2009	347	667K	3K	2K	6K	0.9%	NO
FRANCE-1	France	10/29/2008	488	2806K	10K	5K	11K	0.4%	NO
FRANCE-2	France	11/18/2008	468	4878K	17K	10K	18K	0.4%	NO
FRANCE-3	France	11/5/2008	481	3102K	21K	6K	33K	1.1%	NO
HONGKONG-1	Hong Kong	5/24/2008	646	1229K	10K	6K	3K	0.3%	YES
IRELAND-1	Ireland	5/28/2008	642	11660K	35K	18K	94K	0.8%	YES
ITALY-1	Italy	11/19/2008	467	1076K	4K	3K	2K	0.2%	NO
ITALY-2	Italy	12/5/2008	451	1622K	5K	4K	7K	0.4%	YES
MEXICO-1	Mexico	5/15/2009	290	600K	4K	2K	39K	6.5%	YES
NETHERLANDS-1	Netherlands	5/2/2009	303	1485K	10K	8K	4K	0.3%	YES
NEWZEALAND-1	New Zealand	6/17/2008	622	9528K	39K	12K	295K	3.1%	NO
RUSSIA-1	Russia	2/11/2009	383	13765K	120K	30K	308K	2.2%	NO
SINGAPORE-1	Singapore	3/20/2009	346	514K	2K	2K	1K	0.1%	YES
SPAIN-1	Spain	6/27/2008	612	3017K	11K	5K	28K	0.9%	YES
TURKEY-1	Turkey	6/4/2008	635	8889K	30K	13K	55K	0.6%	YES
UK-1	UK	5/7/2008	663	8124K	24K	13K	152K	1.9%	YES
UK-2	UK	6/27/2008	612	3442K	16K	5K	25K	0.7%	NO
UK-3	UK	2/17/2009	377	494K	6K	4K	5K	1.0%	YES
UK-4	UK	10/5/2008	512	2774K	7K	6K	20K	0.7%	NO
UK-5	UK	6/18/2008	621	433K	4K	3K	1K	0.3%	NO
URUGUAY-1	Uruguay	10/23/2007	860	12297K	46K	10K	79K	0.6%	YES
US-1	US	4/11/2009	324	13484K	57K	35K	486K	3.6%	NO
US-2	US	5/6/2008	664	6309K	14K	10K	35K	0.6%	YES
US-3	US	5/8/2008	662	11868K	29K	15K	262K	2.2%	YES
VENEZUELA-1	Venezuela	5/16/2008	654	10292K	20K	13K	49K	0.5%	NO
Mean			571	5236K	20K	8K	80K	2.9%	
Median			612	3292K	11K	6K	33K	0.8%	

Table 2: Estimation of Hartigan's Dip and Silverman's Tests

	DIP Test (calibrated)		**Silverman's Test**		
	Dip Stat. (lower is unimodal)	Null = 1 mode P-values	Critical Band. (lower is unimodal)	Null = 1 mode P-values	Null $\leq$ 2 modes P-values
ARGENTINA-1	0.07	0.00	1.92	0.00	0.00
ARGENTINA-2	0.03	0.00	1.47	0.00	0.00
AUSTRALIA-1	0.04	0.00	2.21	0.00	0.03
AUSTRALIA-2	0.07	0.04	1.79	0.25	0.24
AUSTRALIA-3	0.02	0.00	1.27	0.00	0.33
AUSTRALIA-4S	0.01	0.00	0.65	0.00	0.00
BRAZIL-1	0.03	0.00	1.12	0.00	0.00
CHILE-1	0.02	0.00	1.74	0.00	0.00
CHINA-1	0.02	0.00	1.34	0.00	0.13
CHINA-2	0.02	0.00	1.61	0.00	0.00
COLOMBIA-1	0.01	0.00	0.66	0.03	0.97
ECUADOR-1	0.02	0.00	1.60	0.00	0.20
FRANCE-1	0.02	0.00	1.40	0.00	0.00
FRANCE-2	0.02	0.00	1.10	0.00	0.01
FRANCE-3	0.04	0.00	0.59	0.00	0.00
HONGKONG-1	0.04	0.00	1.22	0.00	0.00
IRELAND-1	0.05	0.00	1.70	0.00	0.00
ITALY-1	0.02	0.00	0.92	0.03	0.10
ITALY-2	0.03	0.00	1.57	0.00	0.15
MEXICO-1	0.03	0.00	0.81	0.00	0.00
NETHERLANDS-1	0.01	0.00	1.06	0.00	0.07
NEWZEALAND-1	0.05	0.00	2.35	0.00	0.00
RUSSIA-1	0.02	0.00	0.95	0.00	0.00
SINGAPORE-1	0.09	0.00	2.48	0.00	0.33
SPAIN-1	0.03	0.00	1.72	0.00	0.00
TURKEY-1	0.02	0.00	1.62	0.00	0.03
UK-1	0.01	0.00	0.68	0.00	0.00
UK-2	0.01	0.00	0.54	0.00	0.17
UK-3	0.01	0.00	0.83	0.00	0.00
UK-4	0.08	0.00	2.47	0.00	0.00
UK-5	0.03	0.00	0.95	0.06	0.12
URUGUAY-1	0.10	0.00	0.49	0.00	0.00
US-1	0.06	0.00	1.05	0.00	0.00
US-2	0.07	0.00	2.42	0.00	0.01
US-3	0.04	0.00	1.56	0.00	0.00
VENEZUELA-1	0.04	0.00	0.77	0.00	0.00

Table 3: Proportional Mass Test - Distribution centered at 0%

Establishment	Observations	Centered Centered	Point Estimate	Mean of Bootstrap	Standard Deviation	Share above zero
ARGENTINA-1	45946	0	-0.15	-0.15	0.007	0.00
ARGENTINA-2	20283	0	-0.132	-0.133	0.011	0.00
AUSTRALIA-1	9140	0	-0.345	-0.344	0.02	0.00
AUSTRALIA-2	35	0	0	-0.087	0.26	0.41
AUSTRALIA-3	585	0	-0.503	-0.501	0.083	0.00
AUSTRALIA-4	19332	0	0.216	0.216	0.008	1.00
BRAZIL-1	88811	0	-0.092	-0.092	0.005	0.00
CHILE-1	31936	0	-0.236	-0.235	0.01	0.00
CHINA-1	1730	0	-0.241	-0.241	0.039	0.00
CHINA-2	10669	0	-0.62	-0.621	0.021	0.00
COLOMBIA-1	29012	0	-0.011	-0.012	0.007	0.06
ECUADOR-1	1450	0	-0.081	-0.079	0.037	0.01
FRANCE-1	6121	0	-0.171	-0.173	0.019	0.00
FRANCE-2	5309	0	-0.089	-0.088	0.02	0.00
FRANCE-3	20355	0	0.103	0.103	0.008	1.00
HONGKONG-1	933	0	-0.111	-0.113	0.049	0.00
IRELAND-1	18353	0	0.109	0.109	0.008	1.00
ITALY-1	635	0	0.06	0.061	0.052	0.89
ITALY-2	910	0	-0.548	-0.553	0.069	0.00
MEXICO-1	5131	0	0.095	0.094	0.016	1.00
NETHERLANDS-1	2473	0	-0.416	-0.416	0.039	0.00
NEWZEALAND-1	42557	0	-0.293	-0.294	0.008	0.00
RUSSIA-1	70016	0	0.393	0.393	0.004	1.00
SINGAPORE-1	100	0	-1.073	-1.133	0.349	0.00
SPAIN-1	10084	0	-0.196	-0.196	0.016	0.00
TURKEY-1	4597	0	-0.435	-0.435	0.028	0.00
UK-1	71788	0	0.582	0.582	0.003	1.00
UK-2	13597	0	0.638	0.638	0.005	1.00
UK-3	1776	0	0.167	0.168	0.026	1.00
UK-4	1423	0	-1.919	-1.922	0.167	0.00
UK-5	312	0	-0.264	-0.274	0.098	0.00
URUGUAY-1	52454	0	0.959	0.959	0.001	1.00
US-1	210698	0	0.487	0.487	0.002	1.00
US-2	5261	0	-1.192	-1.192	0.05	0.00
US-3	10466	0	0.156	0.156	0.011	1.00
VENEZUELA-1	15779	0	-0.463	-0.463	0.016	0.00

Note: Bootstrap derived from 500 replications. A positive PM score implies unimodality. We reject unimodality if less than 5% of bootstrapped samples have a positive PM score (i.e. when the “Share above Zero” is ¡0.05).

Table 4: Proportional Mass Scores for each quarter
(Distribution centered at 0%)

Estimates through time	Q4.2007	Q1.2008	Q2.2008	Q3.2008	Q4.2008	Q1.2009	Q2.2009	Q3.2009	Q4.2009
ARGENTINA-1	-0.784	-0.641	-0.704	-0.750	-0.675	-0.865	-0.199	0.454	-0.264
ARGENTINA-2	-0.301	-0.239	-0.239	-0.015	-0.010	-0.102	0.103	-0.391	-0.171
AUSTRALIA-1					-0.295	-0.274	-0.106	-0.587	-0.697
AUSTRALIA-2							0.108	0.270	
AUSTRALIA-3					-0.920	-1.403	-0.588	-0.015	-0.134
AUSTRALIA-4			0.234	0.192	0.277				
BRAZIL-1	-1.430	-0.684	-0.459	-0.518	-0.242	-0.513	-0.120	-0.330	0.531
CHILE-1	-0.591	-0.174	-0.263	-0.267	-0.335	-0.222	-0.072	-0.064	0.012
CHINA-1									-0.222
CHINA-2							-1.221		-0.331
COLOMBIA-1	-0.017	0.056	0.003	0.110	-0.019	0.017	-0.077	-0.061	0.055
ECUADOR-1						-0.495	-0.094	-0.347	0.245
FRANCE-1					0.044	-0.338	-0.690	-0.106	-0.219
FRANCE-2					-0.769	-0.235	0.277	-0.358	-0.198
FRANCE-3					-0.600	-0.322	-0.446	0.444	0.357
HONGKONG-1			-0.816	0.029	-1.042	-1.268	-1.316		
IRELAND-1			0.007	0.202	-0.711	-0.117	-0.039	-0.207	-0.549
ITALY-1					-1.195	-0.462	-0.584	0.085	0.009
ITALY-2					0.238	-0.522	-0.470	-1.032	-0.579
MEXICO-1							0.719	-0.529	-0.305
NETHERLANDS-1							-0.740	-0.388	-0.419
NEWZEALAND-1			-0.035	-0.016	-0.315	-0.497	0.012	-0.501	-0.549
RUSSIA-1						-0.190	-0.121	0.624	0.031
SINGAPORE-1							-0.866	-1.238	
SPAIN-1				0.217	0.040	-0.388	-0.392	-0.292	-0.259
TURKEY-1			-0.430	-0.591	-0.438	-0.686	-0.415	-0.247	-0.689
UK-1			0.686	0.689	0.393	0.510	0.658	0.660	0.674
UK-2				0.707	0.605	0.581	0.561	0.374	
UK-3							0.107	0.346	
UK-4						-2.429	-1.132		
UK-5			-0.775	-0.450	-0.379	0.280	0.046		
URUGUAY-1	1.055	1.012	0.906	0.939	0.798	0.907	0.935	0.914	-1.555
US-1							0.479	0.574	0.376
US-2			-1.479	-1.238	-1.800		-1.151	-1.179	-1.114
US-3			-0.487	0.106	-0.190	0.212	0.332	0.491	0.160
VENEZUELA-1			0.508	0.114	0.030	-0.054	-0.901	-0.095	-0.132

Table 5: Comparison with Scanner and CPI sampling methods

	Daily Data	Weekly Average	Monthly Sampling
Mean Dip (Hartigan)	0.035	0.019	0.046
Mean Critical Bandwidth (Silverman)	1.351	0.799	1.471
Mean PM Score	-0.143	0.145	-0.203

Note: Unimodal distributions have lower Dips, lower CBs and positive PMs.

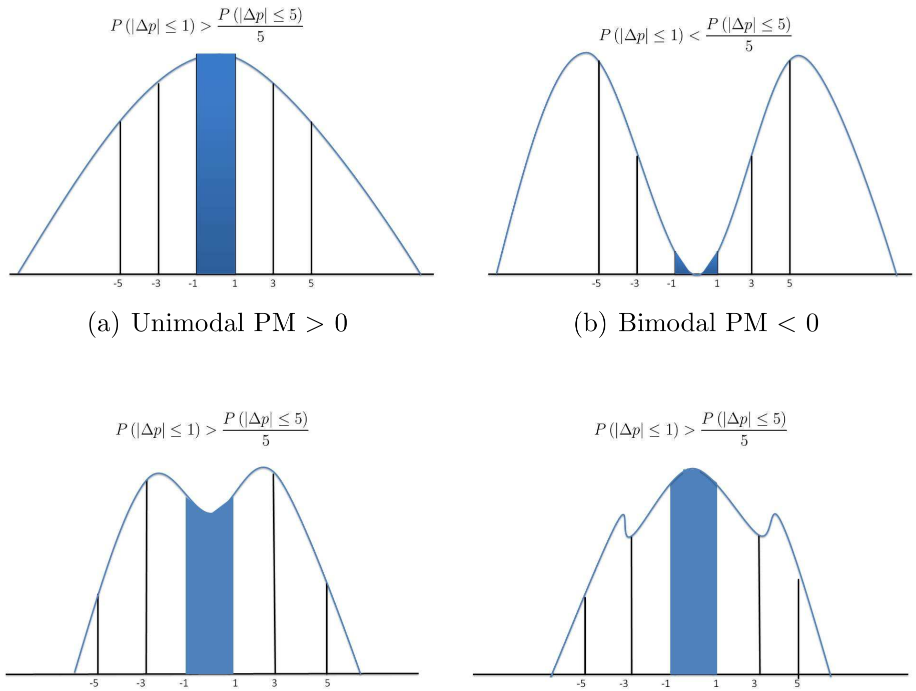
Table 6: PM Score by Supermarket and Categories. Proportion of Categories where Unimodality (centered at zero) is rejected at different levels of significance.

Establishment	cat>0.95	cat>0.90	cat>0.85	cat>0.80	cat>0.75	Cat-Total	Cat-Included
ARGENTINA-1	0.333	0.333	0.333	0.467	0.467	69	15
ARGENTINA-2	0.228	0.228	0.241	0.278	0.304	1419	79
AUSTRALIA-1	0.717	0.739	0.804	0.804	0.848	439	46
AUSTRALIA-2	0.351	0.351	0.364	0.377	0.377	224	77
AUSTRALIA-3	0.769	0.846	0.846	0.846	1.000	167	13
AUSTRALIA-4	0.435	0.478	0.487	0.530	0.557	442	115
BRAZIL-1	0.333	0.333	0.333	0.667	0.667	756	3
CHILE-1	0.081	0.081	0.081	0.084	0.088	1416	297
CHINA-1	0.357	0.401	0.433	0.477	0.545	1346	277
CHINA-2	0.007	0.007	0.007	0.007	0.007	1241	146
ECUADOR-1	1.000	1.000	1.000	1.000	1.000	1093	4
FRANCE-1	0.000	0.000	0.005	0.005	0.005	433	207
FRANCE-2	0.296	0.370	0.370	0.370	0.444	197	27
FRANCE-3	0.700	0.740	0.780	0.800	0.820	143	50
HONGKONG-1	0.493	0.530	0.567	0.593	0.645	687	487
ITALY-1	0.219	0.241	0.257	0.261	0.277	1853	448
ITALY-2	0.333	0.333	0.373	0.373	0.392	675	51
MEXICO-1	0.250	0.250	0.250	0.250	0.250	190	4
NETHERLANDS-1	0.000	0.000	0.000	0.000	0.000	801	29
RUSSIA-1	0.115	0.126	0.130	0.136	0.149	1250	824
SINGAPORE-1	0.770	0.809	0.816	0.816	0.849	293	152
SPAIN-1	0.078	0.109	0.180	0.188	0.242	597	128
URUGUAY-1	0.412	0.412	0.412	0.529	0.588	800	17
US-2	0.714	0.762	0.762	0.762	0.762	411	21
US-3	0.279	0.358	0.380	0.425	0.503	549	179
VENEZUELA-1	0.507	0.587	0.601	0.609	0.630	508	138
Total						19547	3839
Unweighed Mean	0.40	0.42	0.44	0.47	0.50		
Weighted Mean	0.27	0.30	0.31	0.33	0.36		

Table 7: Inflation Dispersion and Rejection of Unimodality

Establishment	Coefficient	Standard Deviation	Tstat
ARGENTINA-1	0.033	0.007	4.780
ARGENTINA-2	0.004	0.002	2.047
AUSTRALIA-1	-0.002	0.006	0.287
AUSTRALIA-4	0.018	0.014	1.273
BRAZIL-1	0.069	0.007	10.599
CHILE-1	0.010	0.019	0.543
CHINA-1	0.029	0.020	1.401
CHINA-2	0.002	0.045	0.046
COLOMBIA-1	0.036	0.016	2.287
ECUADOR-1	-0.031	0.053	0.586
FRANCE-3	0.092	0.042	2.220
IRELAND-1	0.241	0.052	4.671
MEXICO-1	0.167	0.077	2.157
NETHERLANDS-1	0.110	0.050	2.210
NEWZEALAND-1S	0.080	0.020	3.953
RUSSIA-1	0.001	0.014	0.074
SPAIN-1	0.067	0.107	0.630
TURKEY-1	0.041	0.031	1.309
UK-1	0.029	0.003	8.919
UK-2	0.000	0.000	0.000
URUGUAY-1	-0.001	0.003	0.378
US-1	0.008	0.001	6.095
US-3	0.188	0.036	5.200

Figures



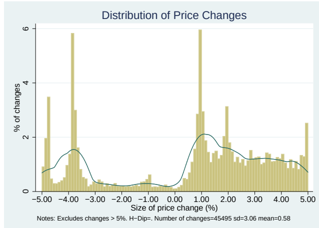
(a) Unimodal PM > 0

(b) Bimodal PM < 0

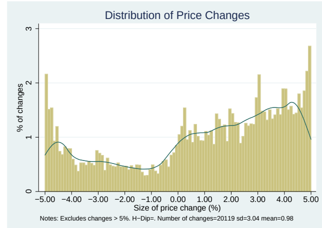
(c) PM > 0

(d) PM > 0

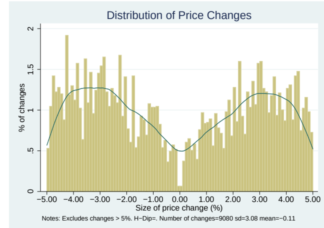
Figure 1: Example of PM values



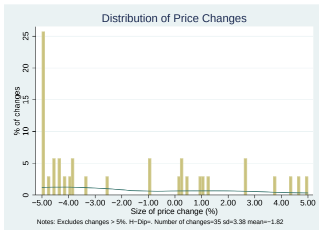
(a) ARGENTINA-1



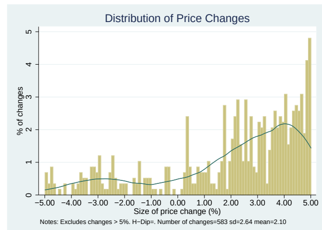
(b) ARGENTINA-2



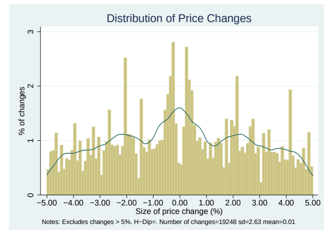
(c) AUSTRALIA-1



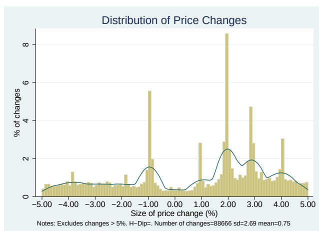
(d) AUSTRALIA-2



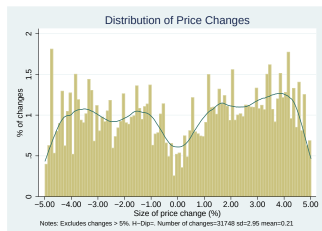
(e) AUSTRALIA-3



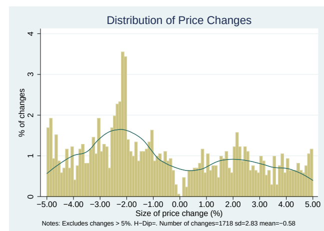
(f) AUSTRALIA-4



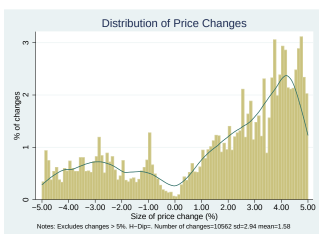
(g) BRAZIL-1



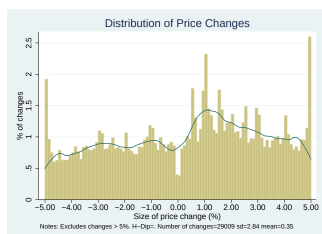
(h) CHILE-1



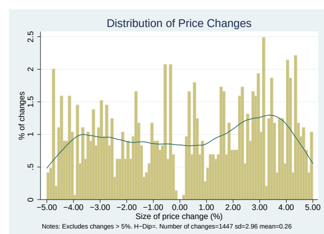
(i) CHINA-1



(j) CHINA-2

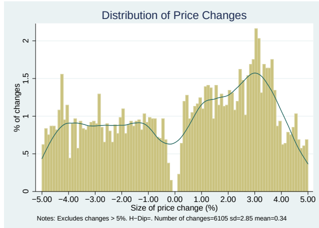


(k) COLOMBIA-1

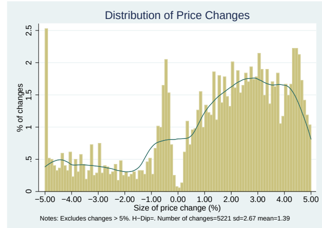


(l) ECUADOR-1

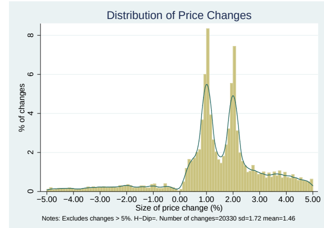
Figure 2: Histogram of Changes - Range -5% to 5%



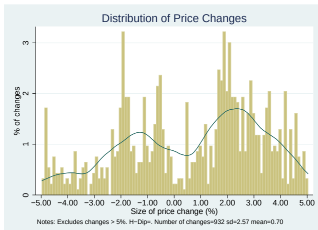
(a) FRANCE-1



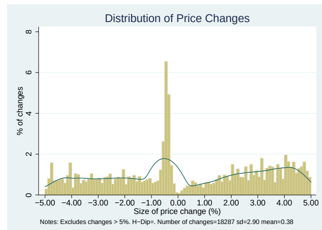
(b) FRANCE-2



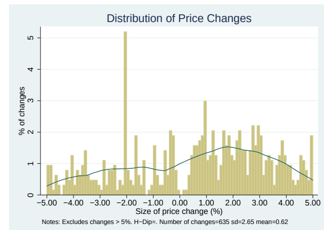
(c) FRANCE-3



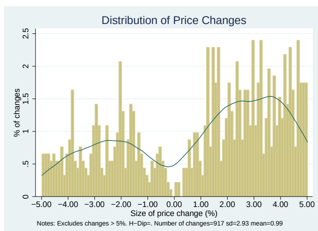
(d) HONGKONG-1



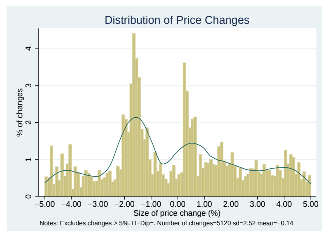
(e) IRELAND-1



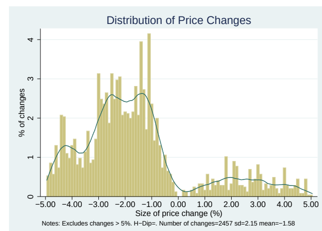
(f) ITALY-1



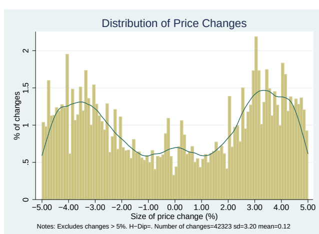
(g) ITALY-2



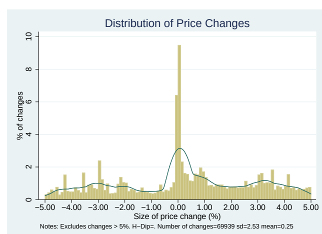
(h) MEXICO-1



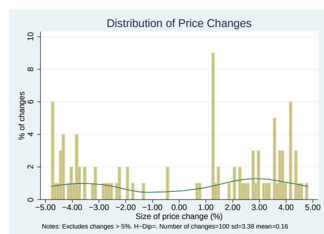
(i) NETHERLANDS-1



(j) NEWZEALAND-1

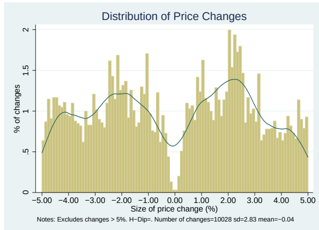


(k) RUSSIA-1

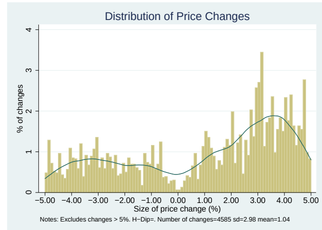


(l) SINGAPORE-1

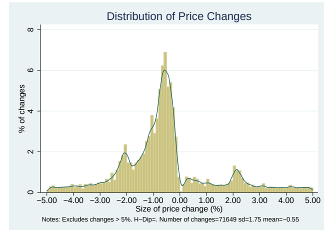
Figure 3: Histogram of Changes - Range -5% to 5%



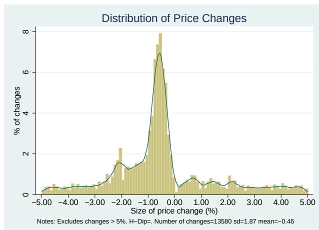
(a) SPAIN-1



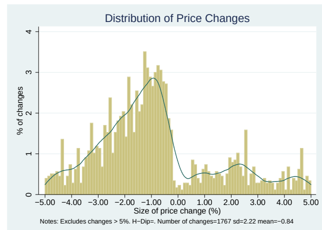
(b) TURKEY-1



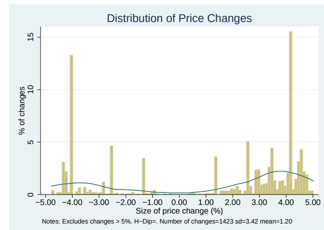
(c) UK-1



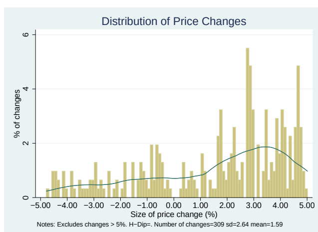
(d) UK-2



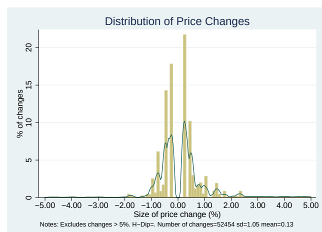
(e) UK-3



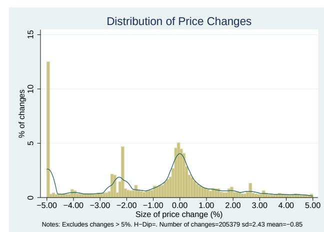
(f) UK-4



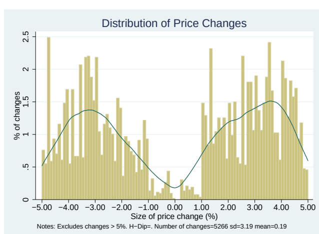
(g) UK-5



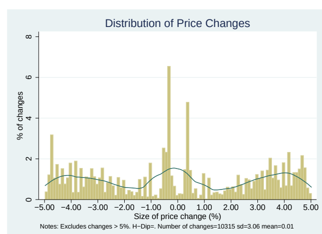
(h) URUGUAY-1



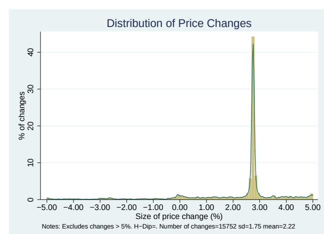
(i) US-1



(j) US-2

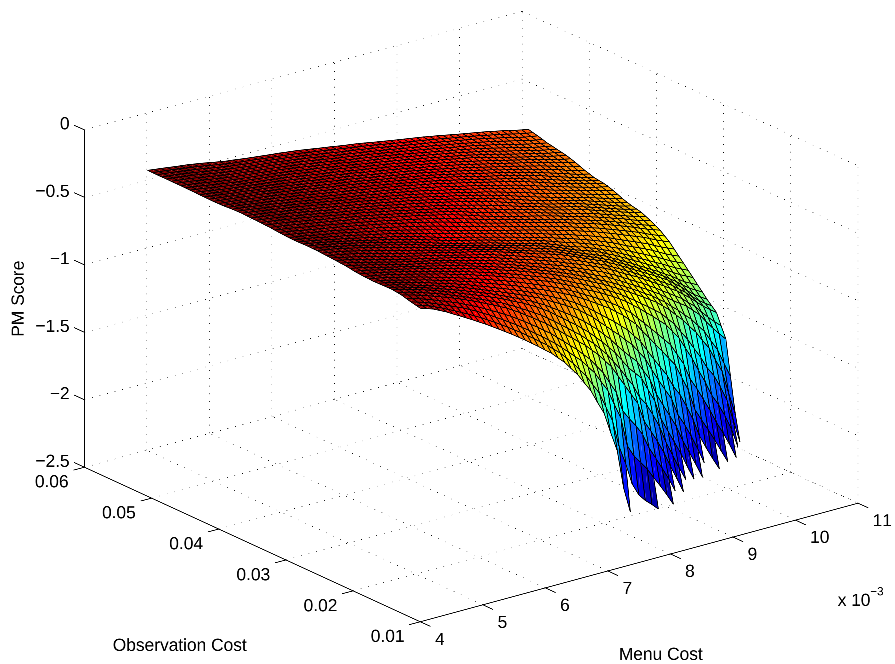


(k) US-3

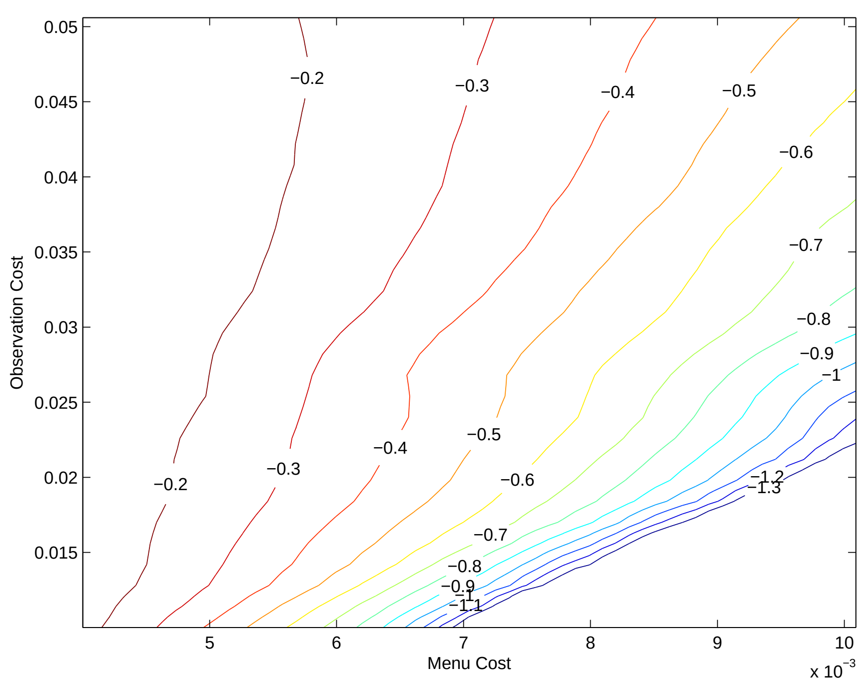


(l) VENEZUELA-1

Figure 4: Histogram of Changes - Range -5% to 5%



(a) Proportional Mass Score



(b) Proportional Mass Contours

Figure 5: PM Score and Contours for 0-0.05 percent range

–APPENDIX–

The Distribution of the Size of Price Changes

Alberto Cavallo, Harvard University
Roberto Rigobon, MIT & NBER

May 27, 2011

A Appendix

A.1 Additional Tables and Figures

Table A1: Implied Mean and Median Durations

	Mean (days)	Median (days)
ARGENTINA-1	122	82
ARGENTINA-2	73	50
AUSTRALIA-1	16	11
AUSTRALIA-2	139	67
AUSTRALIA-3	465	526
AUSTRALIA-4	42	29
BRAZIL-1	70	53
CHILE-1	171	96
CHINA-1	165	112
CHINA-2	188	163
COLOMBIA-1	85	57
ECUADOR-1	136	98
FRANCE-1	260	251
FRANCE-2	251	238
FRANCE-3	96	78
HONGKONG-1	315	291
IRELAND-1	144	101
ITALY-1	374	479
ITALY-2	243	236
MEXICO-1	48	29
NETHERLANDS-1	175	140
NEWZEALAND-1	49	23
RUSSIA-1	66	55
SINGAPORE-1	244	229
SPAIN-1	106	81
TURKEY-1	196	126
UK-1	106	61
UK-2	113	78
UK-3	125	87
UK-4	386	470
UK-5	105	53
URUGUAY-1	173	131
US-1	66	28
US-2	175	109
US-3	89	45
VENEZUELA-1	226	164
Mean	159	135
Median	136	87

Note: Implied Durations using method in Bils and Klenow (2004)

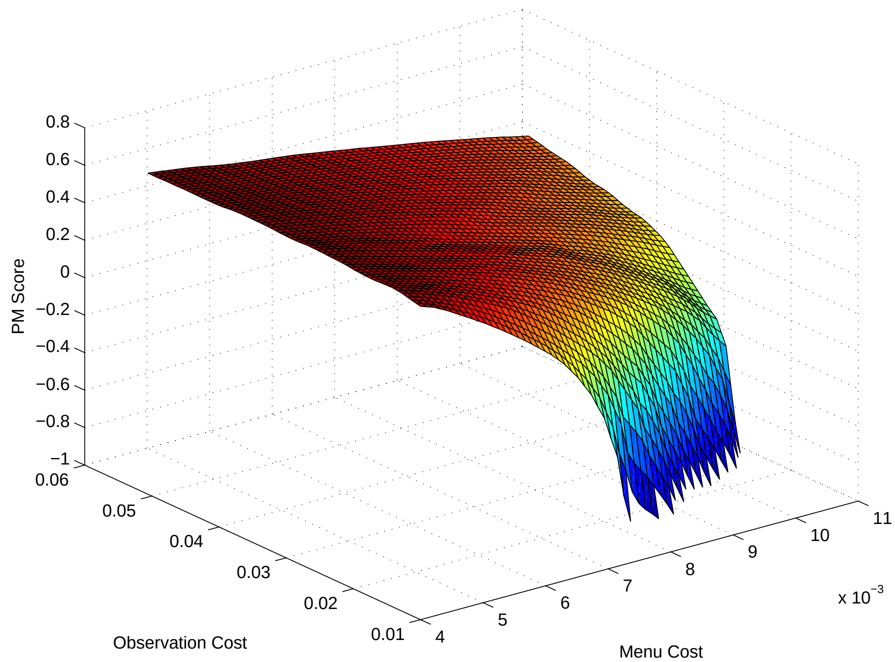
Table A2: Proportional Mass Test - Distribution centered at the Largest Mode

Establishment	Observations	Centered Centered	Point Estimate	Mean of Bootstrap	Standard Deviation	Share above zero
ARGENTINA-1	50598	0.9	0.178	0.178	0.005	1.00
ARGENTINA-2	36321	4.9	0.287	0.287	0.005	1.00
AUSTRALIA-1	9606	-2.9	0.109	0.109	0.012	1.00
AUSTRALIA-2	105	-3.7	-0.565	-0.597	0.232	0.00
AUSTRALIA-3	1078	4.9	0.225	0.221	0.031	1.00
AUSTRALIA-4	19330	-0.3	0.213	0.213	0.008	1.00
BRAZIL-1	90226	1.9	0.38	0.38	0.003	1.00
CHILE-1	32092	4.1	0.129	0.128	0.007	1.00
CHINA-1	1804	-2.1	0.285	0.284	0.023	1.00
CHINA-2	12691	4.1	0.432	0.432	0.008	1.00
COLOMBIA-1	29651	1.1	0.159	0.159	0.007	1.00
ECUADOR-1	1552	4.1	0.115	0.114	0.033	1.00
FRANCE-1	5486	3.1	0.313	0.313	0.013	1.00
FRANCE-2	6314	4.5	0.305	0.303	0.013	1.00
FRANCE-3	20869	1.1	0.725	0.725	0.004	1.00
HONGKONG-1	1086	1.9	0.22	0.217	0.033	1.00
IRELAND-1	17991	-0.5	0.152	0.153	0.008	1.00
ITALY-1	594	-2.1	-0.095	-0.091	0.058	0.05
ITALY-2	1020	1.5	0.077	0.075	0.039	0.97
MEXICO-1	5021	-1.7	0.342	0.341	0.013	1.00
NETHERLANDS-1	2603	-1.1	0.381	0.381	0.017	1.00
NEWZEALAND-1	46034	3.1	0.127	0.127	0.005	1.00
RUSSIA-1	70366	0.1	0.406	0.406	0.003	1.00
SINGAPORE-1	83	1.3	0.09	0.085	0.129	0.74
SPAIN-1	9791	2.1	0.177	0.177	0.012	1.00
TURKEY-1	6437	3.1	0.118	0.118	0.014	1.00
UK-1	71839	-0.5	0.711	0.711	0.002	1.00
UK-2	13554	-0.5	0.695	0.695	0.005	1.00
UK-3	1769	-0.7	0.46	0.459	0.019	1.00
UK-4	2132	4.1	0.31	0.31	0.021	1.00
UK-5	425	2.9	0.153	0.153	0.058	0.99
URUGUAY-1	52651	0.3	0.908	0.908	0.001	1.00
US-1	230505	-0.1	0.423	0.423	0.002	1.00
US-2	4684	-3.3	0.261	0.261	0.016	1.00
US-3	10780	-0.3	0.159	0.16	0.011	1.00
VENEZUELA-1	18146	2.7	0.735	0.735	0.004	1.00

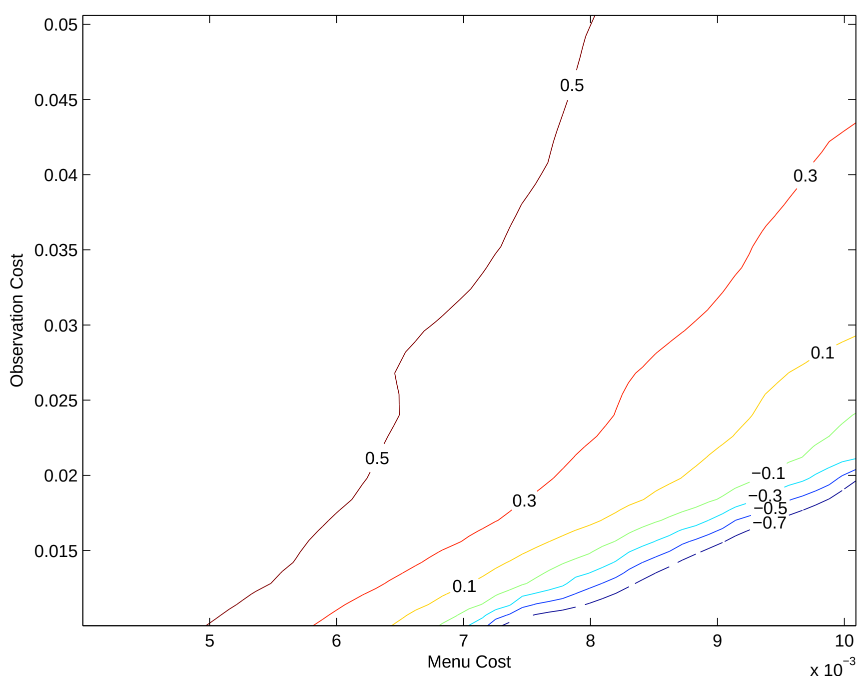
Note: Bootstrap derived from 500 replications.

Table A3: Share of bootstraps with unimodal PM scores in each quarter, when distribution is centered at the largest mode.

Estimates through time	Q4.2007	Q1.2008	Q2.2008	Q3.2008	Q4.2008	Q1.2009	Q2.2009	Q3.2009	Q4.2009
ARGENTINA-1	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
ARGENTINA-2	1.00	1.00	1.00	1.00	1.00	0.98	0.80	1.00	0.78
AUSTRALIA-1					0.92	1.00	0.02	1.00	1.00
AUSTRALIA-2							0.10	0.86	
AUSTRALIA-3					1.00	1.00	0.95	0.41	0.51
AUSTRALIA-4			1.00	1.00	1.00				
BRAZIL-1	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
CHILE-1	1.00	1.00	1.00	1.00	1.00	1.00	0.07	0.04	1.00
CHINA-1									1.00
CHINA-2							1.00		1.00
COLOMBIA-1	0.98	0.93	1.00	0.01	1.00	1.00	1.00	1.00	1.00
ECUADOR-1						0.95	0.95	1.00	1.00
FRANCE-1					1.00	1.00	1.00	1.00	0.88
FRANCE-2					1.00	1.00	1.00	1.00	0.01
FRANCE-3					1.00	1.00	1.00	1.00	1.00
HONGKONG-1			1.00	1.00	1.00	0.93	0.75	1.00	1.00
IRELAND-1			0.97	1.00	0.98	1.00	0.96	1.00	1.00
ITALY-1					0.65	0.94	0.99	0.89	0.99
ITALY-2					0.81	1.00	0.98	0.88	0.99
MEXICO-1							1.00	1.00	1.00
NETHERLANDS-1							1.00	1.00	1.00
NEWZEALAND-1			1.00	0.31	1.00	1.00	0.71	1.00	1.00
RUSSIA-1						1.00	1.00	1.00	1.00
SINGAPORE-1							0.88	0.98	0.99
SPAIN-1				1.00	0.99	0.20	1.00	1.00	1.00
TURKEY-1			0.79	1.00	1.00	0.99	1.00	1.00	1.00
UK-1			1.00	1.00	1.00	1.00	1.00	1.00	1.00
UK-2			1.00	1.00	1.00	1.00	1.00	1.00	
UK-3							1.00	1.00	
UK-4					1.00	0.91	1.00	1.00	0.99
UK-5			0.85	1.00	0.98	0.82	0.81		
URUGUAY-1	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
US-1							1.00	1.00	1.00
US-2			1.00	1.00	1.00		1.00	1.00	1.00
US-3			1.00	1.00	0.00	1.00	1.00	1.00	1.00
VENEZUELA-1			1.00	1.00	0.52	1.00	1.00	1.00	0.88
Supermarkets with some “bimodality”	1	1	4	3	8	10	14	6	9
Supermarkets	6	6	17	18	26	26	34	32	31
Ratio with “Bimodality”	0.17	0.17	0.24	0.17	0.31	0.38	0.41	0.19	0.29

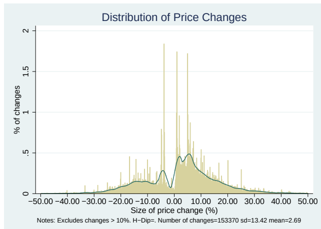


(a) Proportional Mass Score

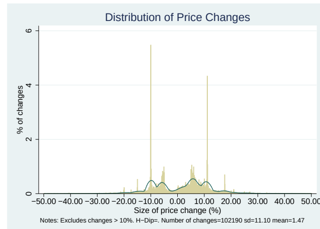


(b) Proportional Mass Contours

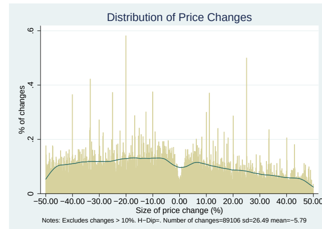
Figure A1: PM Score and Contours for 0-0.20 percent range



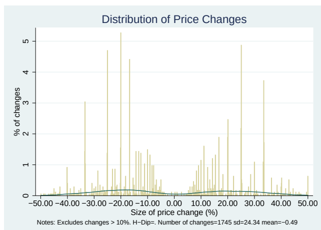
(a) ARGENTINA-1



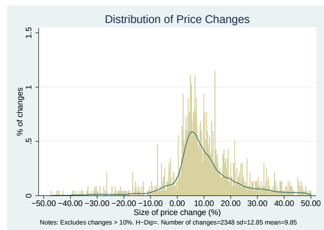
(b) ARGENTINA-2



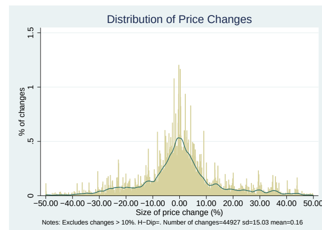
(c) AUSTRALIA-1



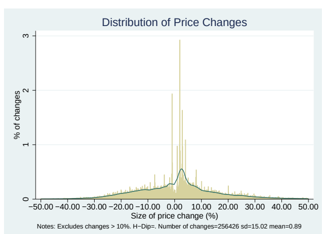
(d) AUSTRALIA-2



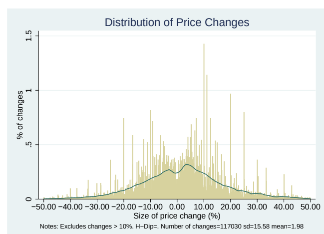
(e) AUSTRALIA-3



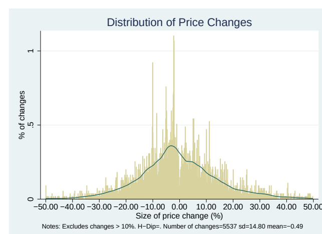
(f) AUSTRALIA-4



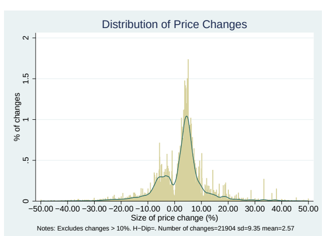
(g) BRAZIL-1



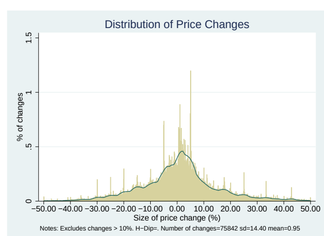
(h) CHILE-1



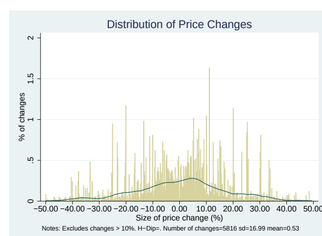
(i) CHINA-1



(j) CHINA-2

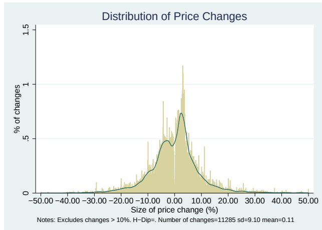


(k) COLOMBIA-1

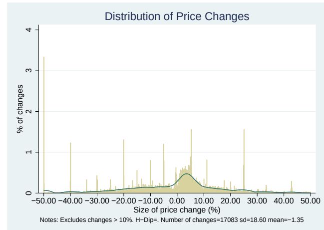


(l) ECUADOR-1

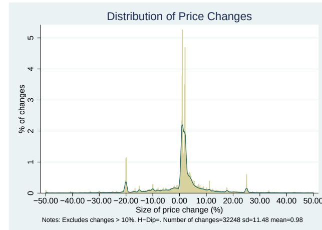
Figure A2: Histogram of Changes - Range -50% to 50%



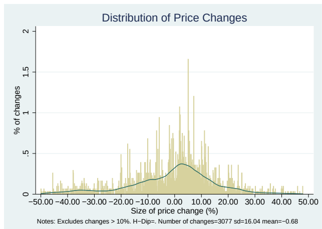
(a) FRANCE-1



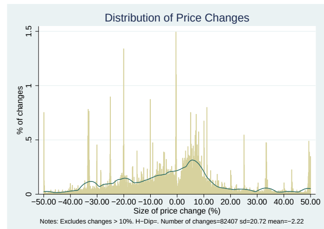
(b) FRANCE-2



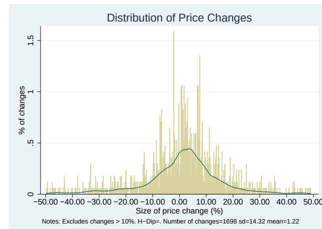
(c) FRANCE-3



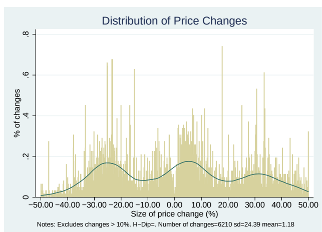
(d) HONGKONG-1



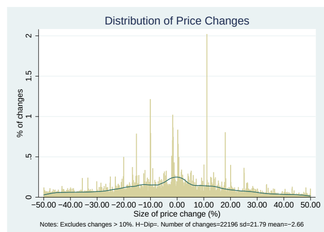
(e) IRELAND-1



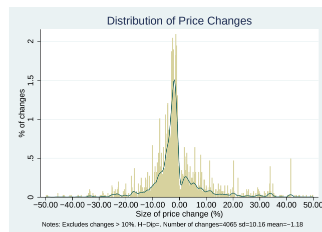
(f) ITALY-1



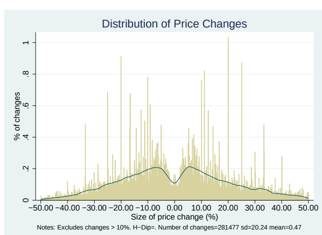
(g) ITALY-2



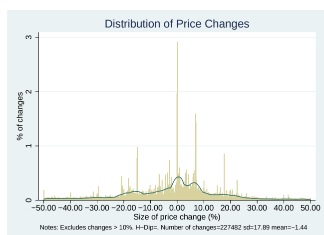
(h) MEXICO-1



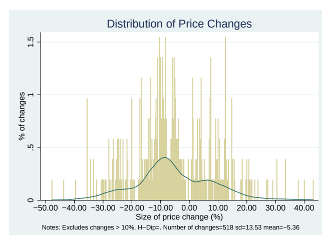
(i) NETHERLANDS-1



(j) NEWZEALAND-1

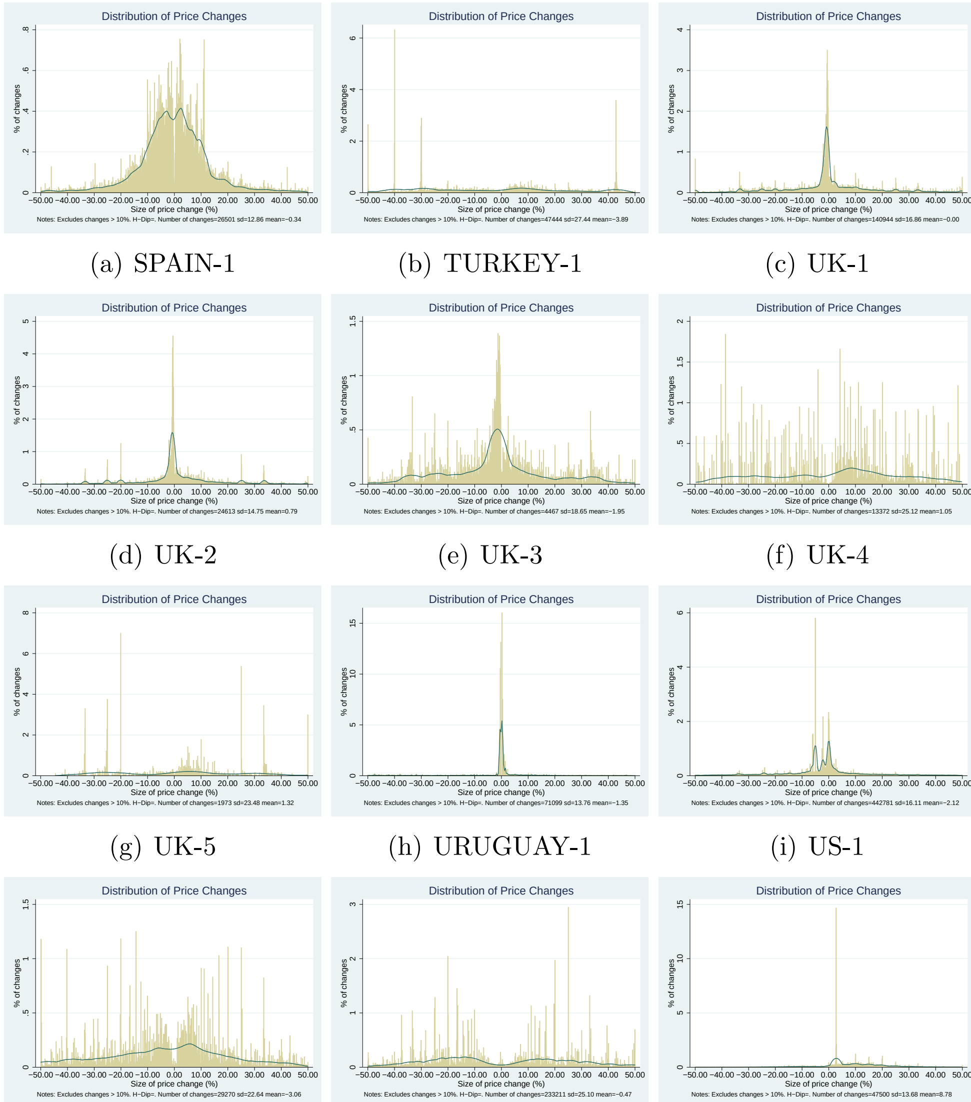


(k) RUSSIA-1



(l) SINGAPORE-1

Figure A3: Histogram of Changes - Range -50% to 50%



(a) SPAIN-1 (b) TURKEY-1 (c) UK-1

(d) UK-2 (e) UK-3 (f) UK-4

(g) UK-5 (h) URUGUAY-1 (i) US-1

(j) US-2 (k) US-3 (l) VENEZUELA-1

Figure A4: Histogram of Changes - Range -50% to 50%