EIEF WORKING PAPER SERIES

Einaudi Institute for Economics and Finance

EIEF Working Paper 16/08
April 2016

The real effects of monetary shocks in sticky price models: a sufficient statistic approach

by

Fernando Alvarez
(University of Chicago)
Hervé Le Bihan
(Banque de France)
Francesco Lippi
(EIEF and University of Sassari)

The real effects of monetary shocks in sticky price models: a sufficient statistic approach*

Fernando Alvarez
University of Chicago

Hervé Le Bihan
Banque de France

Francesco Lippi
EIEF & University of Sassari

April 27, 2016

Abstract

We prove that the ratio of kurtosis to the frequency of price changes is a sufficient statistic for the real effects of monetary shocks, measured by the cumulated output response following the shock. The sufficient statistic result holds in a large class of models which includes Taylor (1980), Calvo (1983), Reis (2006), Golosov and Lucas (2007), Nakamura and Steinsson (2010), Midrigan (2011) and Alvarez and Lippi (2014). Several models in this class are able to account for the positive excess kurtosis of the size-distribution of price changes that appears in the data. We review empirical measures of kurtosis and frequency and conclude that a model that successfully matches the micro evidence on kurtosis and frequency produces real effects that are about 4 times larger than in the Golosov-Lucas model, and about 30% below those of the Calvo model. We discuss the robustness of our results to changes in the setup, including small inflation and leptokurtic cost shocks.

JEL Classification Numbers: E3, E5

Key Words: price setting, menu cost, Calvo pricing, micro evidence, kurtosis of price changes, sufficient statistic, output response to monetary shocks

*We benefited from the comments of seminar participants at the 2014 NBER EF&G meeting in NY, at the Chicago Fed, Federal Reserve Board, Minneapolis Fed, and the following universities: Bocconi, Bologna, Firenze, MIT, Ohio State, HECER and Lausanne. We are grateful to Alberto Cavallo, Pete Klenow, Oleksiy Kryvtsov, Joseph Vavra and Fredrik Wulfsberg for providing us with several statistics not available in their papers. We thank David Argente and for proving us with evidence for the US scanner data. We thank the INSEE and the SymphonyIRI Group, Inc. for making the data available. All estimates and analysis in this paper, based on data provided by SymphonyIRI Group, Inc. are by the authors and not by SymphonyIRI Group, Inc. We thank the Fondation Banque de France for supporting this project. The views expressed in the paper are of the authors and do not necessarily reflect those of the Banque de France. Part of the research for this paper was sponsored by the ERC advanced grant 324008. Roberto Robatto, David Argente and Jean Flemming provided excellent research assistance. A previous version of the paper was titled "Small and large price changes and the propagation of monetary shocks".

1 Introduction

This paper provides an analytical characterization of the steady state cross-sectional moments, and of the response of output to an unexpected monetary shock, in models with sticky prices. By combining the assumptions of multiproduct firms and random menu costs the model is able to produce, in various degrees, both the small and large price changes that have been documented in the micro data starting with Kashyap (1995). Different sticky price set-ups, spanning the models of Taylor (1980), Calvo (1983), Golosov and Lucas (2007), some versions of the "CalvoPlus" model by Nakamura and Steinsson (2010), the "rational inattentiveness" model by Reis (2006), as well as the multi-product models of Midrigan (2011), Bhattarai and Schoenle (2014) and Alvarez and Lippi (2014), are nested by our model. This unified framework allows us to unveil which assumptions are required to obtain each model as an optimal mechanism.

The main analytical result of the paper is that, in a large class of models that includes those listed above, the total cumulative output effect of a small unexpected monetary shock depends on the ratio between two steady-state statistics: the kurtosis of the size-distribution of price changes $Kur(\Delta p_i)$ and the average number of price changes per year $N(\Delta p_i)$. Formally, given the labor supply elasticity $1/\epsilon - 1$ and a small monetary shock δ, we show that the cumulative output $\mathcal{M}$, namely the area under the output impulse response function, is

$$\mathcal{M} = \frac{\delta}{6\,\epsilon}\,\frac{Kur(\Delta p_i)}{N(\Delta p_i)} \quad . \tag{1}$$

The impact of the frequency $N\left(\Delta p_i\right)$ on the real output effect is understood in the literature and motivates a large body of empirical literature. The main novelty is that the effect of $Kur\left(\Delta p_i\right)$ is equally important, and motivates our interest to discuss its measurement and report new evidence on it. For a symmetric distribution, kurtosis is a scale-free statistic describing its shape, specifically its peakedness: the extent to which "large" and "small" observations (in absolute value) appear relative to intermediate values. We show that this statistic embodies the extent to which "selection" of price changes occurs. The selection effect, a terminology introduced by Golosov and Lucas (2007), indicates that the firms that change prices after the monetary shock are the firms whose prices are in greatest need of adjustment, not a random sample. Selection gives rise to large price adjustments after the shock, so that the CPI response is fast. Such selection is absent in Calvo where the adjusting firms are randomly chosen and, after a shock, the size of the average price change (across adjusting firms) is constant, so that the CPI rises more slowly and the real effects are more persistent. Surprisingly, the kurtosis of the steady-state distribution of the size of price changes fully encodes the selection effect. Intuitively, in the Golosov-Lucas model (steady state) price changes are concentrated around two values: very large and very small, which imply the smallest value of kurtosis (equal to one). In contrast, the size distribution of price adjustments under a Calvo mechanism is very peaked, featuring a large mass of very small as well as some very large price changes, which is exactly what is captured by the high kurtosis predicted by the Calvo model (equal to six).

In addition to selection in the *size* of price changes, recent contributions have highlighted a related selection effect in the *timing* of price changes, see e.g. Kiley (2002); Sheedy (2010); Carvalho and Schwartzman (2015); Alvarez, Lippi, and Paciello (2016). This paper shows that the selection concerning the timing is also encoded in the kurtosis of price changes. For instance, in the models of Taylor and Calvo, calibrated to the same mean frequency of price changes $N(\Delta p_i)$, the *size* of the average price change across adjusting firms is constant (after a monetary shock), so there is no selection concerning the size. Yet the real cumulative output effect in Calvo is twice the effect in Taylor. This happens because in Taylor the time elapsed between adjustments is a constant $T = 1/N(\Delta p_i)$, while in Calvo it has an exponential distribution (with mean T), with a thick right tail of firms that adjust very late. Notice how these features are captured by kurtosis: in Taylor the constant time between adjustments T implies that price changes are drawn from a normal distribution, hence kurtosis is three. In Calvo, instead, the exponential distribution of adjustment times implies that price changes are drawn from a mixture of normals with different variances, and hence a higher kurtosis (equal to six).

The main advantage of our theoretical result is its robustness: equation (1) allows us to discuss the output effect of monetary shocks in a large class of models without having to solve for the whole general equilibrium or provide details about several other modeling choices. We see this result in the spirit of the sufficient statistic approach introduced in the public finance literature by Chetty (2009) and applied to the new trade literature by Arkolakis, Costinot, and Rodriguez-Clare (2012): the identification of a robust relationship that contains useful economic information, independently of many details of the model, and which can be measured in the data. A key assumption for equation (1) to hold is that the distribution of the cost shocks faced by the firms is normal and that inflation is small. In Section 5 we discuss the realism of these assumptions and explore the robustness of our model to the assumption of non-normal cost shocks that match actual cost data. We show that equation (1) remains accurate in predicting the real effects of monetary shocks.

The paper is organized as follows: we conclude the introduction with an overview of the main results and a brief summary of the related literature. Section 2 discusses the measurement of the kurtosis of price changes, a central statistic in our theory. Section 3 presents the theoretical model and its cross section predictions. Section 4 characterizes analytically the effect of an unexpected monetary shock. Section 5 reviews the scope of our main result and its applicability to actual economies. Section 6 concludes by discussing the robustness of our main result to settings where price stickiness originates from information frictions (e.g. rational inattentiveness) rather than from a menu cost.

Overview of main results

Since kurtosis is a sufficient statistic for the real effect of monetary policy in the class of models we analyze, we begin by discussing its estimation. We identify two potential sources of upward bias: heterogeneity (across types of goods and outlets) and small measurement errors (incorrectly imputing small price changes when there was none). We then analyze the size distribution of price changes using two datasets, the French CPI and US Dominick's data. We show that, after controlling for cross-sectional heterogeneity and correcting for measurement error, measured kurtosis is in the vicinity of 4 in a large a sample of low-inflation countries.

We develop an analytical model that features both the small and large price changes which lead to excess kurtosis. The model extends the multi-product setup developed in Alvarez and Lippi (2014), where the fixed menu cost applies to a bundle of n goods sold by each firm. Each good is subject to idiosyncratic cost shocks that create a motive for price adjustment. The shocks are uncorrelated across goods and we assume zero inflation (both assumptions can be relaxed: we show analytically in Section 4 that the results for a small inflation rate are virtually identical to the ones for zero inflation). The multi-product assumption generates the extreme price changes, both small and large. We extend that setup by introducing random menu costs, a feature that produces a positive excess kurtosis of the size-distribution of price changes.[1] In particular, we assume that at an exogenous rate λ each firm receives an opportunity to adjust its price at no cost, as in a Calvo setup. The model has four fundamental parameters: the size of the menu cost relative to curvature of the profit function ψ/B, the volatility of idiosyncratic cost shocks σ^2, the number of products n and the arrival rate of free adjustments λ.

The model yields three new theoretical results. First we characterize how the inaction set behaves as a function of the parameters. For a small menu cost ψ/B the model behaves as in Barro (1972), Dixit (1991) and Golosov and Lucas (2007): the size of the inaction set displays the usual high sensitivity (i.e. a "quartic root") with respect to the cost and the volatility of the shocks σ^2 (the option value effect). Interestingly, the decision rule is unaffected by the presence of the free adjustments as long as the menu cost is small. The decision rule changes

[1] Absent random menu cost the multi-product model can at most produce a zero excess kurtosis.

substantially for large menu costs, an assumption that is useful to generate behavior that approaches that of the Calvo model. In this case the size of the inaction set changes with the square root of the menu cost and the arrival rate, and somewhat surprisingly it becomes unresponsive to the volatility of idiosyncratic shock σ^2, so that changes in the uncertainty faced by firms induce no change in behavior (i.e. there is no option value).

Second, by aggregating the optimal decision rules across firms we characterize the frequency $N(\Delta p_i)$, standard deviation $Std(\Delta p_i)$, and shape of the distribution of the price changes, e.g. its kurtosis: $Kur(\Delta p_i)$. We show that for any pair of parameters $\{n, \lambda\}$, the two remaining parameters $\{\psi/B, \sigma\}$ map one-to-one onto the observables $N(\Delta p_i), Std(\Delta p_i)$. This mapping is convenient for the analysis because it allows us to "freeze" the two observables $N(\Delta p_i)$ and $Std(\Delta p_i)$, which one can take from the data, while retaining the flexibility to accommodate various shapes for the size-distribution of price changes as well as various data on the cost of price adjustment. In particular, we show that the shape of the distribution of price changes can be written exclusively in terms of n and the fraction of free-adjustments $\ell \equiv \lambda/N(\Delta p_i)$. In our model the shape of the distribution of price changes ranges from bimodal (for the model where $\ell = 0$ and $n = 1$ as in the Golosov-Lucas model) to Normal (for $n = \infty$ and $\ell = 0$, our version of Taylor's model), and Laplace (in the case $\ell = 1$ for any n, our version of the Calvo model). In those three models the kurtosis of price changes is, respectively, 1, 3 and 6. In our set-up a given kurtosis may be obtained by different combinations of n and ℓ, yet we argue that models with high n yield a better representation of the cross-sectional data because it eliminates the predominant mass of large price changes that arises in models where n is small.[2] The model is thus able to match several cross sectional features of the micro data, such as the frequency, standard deviation and kurtosis of price changes, using realistic small values for the menu-cost of price adjustments.

Third, we use the model to characterize analytically the impulse response of the aggregate economy to a once-and-for-all unexpected permanent increase of the money supply in Section 4. The effect of a monetary shock depends on the shock size. Large shocks (relative to the size of price adjustments) lead to almost all firms adjusting prices and hence imply neutrality. We fully characterize the minimum size of the shock that delivers this neutrality. Instead, small shocks, such as those found in empirical impulse responses, yield real output effects whose cumulative effect is completely encoded by the frequency and kurtosis of price changes. This result is new in the literature and was discussed in equation (1). To further illustrate our result, we also present the impulse response function associated with the total cumulative output effect. We compute this impulse response by a calibration which shows that a model that successfully matches the cross-sectional micro evidence produces persistent

[2] We show that kurtosis is increasing in both n and ℓ (i.e. both partial derivatives are positive).

real effects that are not too different from what is seen in several empirical studies of the propagation of monetary shocks. In particular, we find that the half life of a monetary shock is about 4 times more persistent than in the Golosov-Lucas model, but 30% less than in the Calvo model.

While the main model we present assumes ex-ante identical firms, in Appendix E we present an extension that allows for heterogeneity in the frequency and the kurtosis of price changes across different sectors of the economy. It is clear from equation (1) that this heterogeneity is the only one that matters for the real effects of monetary shocks. Other forms of heterogeneity, such as sector specific menu costs, might give rise to different standard deviations of price changes which are irrelevant for the monetary transmission (at the sector level) but might bias the economy-wide measurement of kurtosis if heterogeneity is not accounted for, as discussed in Section 2.

Other related literature

In addition to the papers cited above, our analysis relates to a large literature on the propagation of monetary shocks in sticky price models, unifying earlier results that compare the propagation in the Calvo model with the propagation in either the Taylor or the menu cost model of sticky prices. Kiley (2002) showed that, controlling for the frequency of price changes, the response of output is more persistent under Calvo than under Taylor contracts. Golosov and Lucas (2007) compared a menu cost and a Calvo model, with the same frequency of price changes, and find that the half-life of the response to the shock in Calvo is about five times larger than in the menu cost model. Our model is related to the CalvoPlus model of Nakamura and Steinsson (2010) who consider firms facing idiosyncratic cost shocks as well as a menu cost that oscillates randomly between a large and a small value. In their model, like in ours, the random menu cost makes the adjustment decision state dependent, a feature that dampens substantially the real effect of monetary shocks relative to the Calvo model. Our paper also relates to the random menu cost model with idiosyncratic shocks by Dotsey, King, and Wolman (2009), as well as to Vavra (2014), who studies the propagation of shocks in the presence of aggregate volatility shocks. Given the numerical nature of these contributions, these papers do not provide an explicit map between the model fundamentals, the steady state statistics and the propagation of shocks. We see our results as complementary to these numerical analyses. Our model allows for an analytical characterization of the firm's decision rule, the economy's steady state statistics, the identification of the key model parameters, as well as a characterization of the relationship between these statistics and the size of the output effect of monetary shocks. More recently, Karadi and Reiff (2014) have explored menu cost models where the innovation to cost shocks are assumed to be leptokurtic instead of normal. Their model nests special cases, such as the model with normal cost shocks, as well as models a la Midrigan (2011) or Gertler and Leahy (2008) where shocks are either large or zero (and thus cost shocks have a large kurtosis), and a continuum of intermediate cases. They show that leptokurtic cost shocks may boost the output effect of monetary shocks in principle, even though in their calibration to actual cost data they find near money neutrality. While our benchmark model features normal cost shocks, Section 5 uses a data-consistent parametrization to explore the consequences of leptokurtic cost shocks on our main results. We show that equation (1) continues to convey a reliable approximation for the real effects of monetary shocks.

2 Measuring the kurtosis of price changes

A vast amount of research has investigated the size of price changes at the microeconomic level in the past decade (e.g. Bils and Klenow (2004), Nakamura and Steinsson (2008), Klenow and Malin (2010)). A robust empirical pattern is that the size distribution of price changes exhibits a large amount of small, as well as large, price changes. This feature of the distribution is reflected in a kurtosis that is above the one of the normal distribution (i.e. larger than 3). This pattern does not only reflect cross sectional differences between goods types, it also appears at the very disaggregated product level, i.e. a given good typically records both small and large price changes (see Kashyap (1995) for an early documentation of this fact). Most theoretical models fail to produce a size-distribution for price changes with such features. The model we present in the next section will be able to match such patterns and will assign kurtosis a central role in the transmission of monetary shocks. Because kurtosis is a sufficient statistic for the real effect of monetary policy in the class of models we analyze, we discuss two important sources of upward bias that arise in estimation: heterogeneity (across types of goods and outlets) and measurement errors (incorrectly imputing small price changes when there was none). We then analyze the size distribution of price changes using two datasets, the French CPI and US Dominick's data. We find that, after controlling for cross-sectional heterogeneity and correcting for measurement error, the kurtosis is substantially smaller than what is measured in the raw data. We conclude with an overview of the existing evidence from other sources, which shows that the value of kurtosis is in the vicinity of 4 in a large a sample of low-inflation countries. In other words, the distribution lays between a Normal and a Laplace distribution. The other key sufficient statistic for the real effects of monetary policy in our model is the average frequency of price changes, a feature shared with many models. We do not explore the measurement of the frequency in detail since this is the subject of many papers. We do however highlight that measurement error also affects the estimates of price durations, leading to overestimating the frequency of price changes, as documented recently by Cavallo (2015) on US data.

2.1 Accounting for measurement error and heterogeneity

Two concerns arise about the measurement of kurtosis using the micro data: heterogeneity and measurement errors. Heterogeneity refers to the fact that price data in general combine a wide variety of goods. A well–known statistical result is that a mixture of distributions with different variances and the same kurtosis has a larger kurtosis than each sub-population. Likewise, measurement error may bias the estimation of kurtosis as well as of the frequency of price changes. We use a simple statistical model to illustrate these points. Let the observed price changes Δp_m be given by a mixture of three independent zero-mean distributions, with a common kurtosis k and standard deviations σ_j. In particular, let $\Delta p_m = \mathcal{I}_1 \Delta p_1 + \mathcal{I}_2 \Delta p_2 + \mathcal{I}_e \Delta p_e$ where $\mathcal{I}_j$ is an indicator variable that refers to the distribution that originates the price change $j = \{1, 2, e\}$. The distribution indexed by $j = e$ describes observations that are spurious, i.e. due to measurement error. The distributions indexed by $j = \{1, 2\}$ refer to two subpopulations of goods, with standard deviations σ_1 and σ_2. Let $1-\zeta$ be the probability that the observation is a measurement error, and π_1 the probability that the price change is drawn from population 1 (conditional on a true price change being observed). The objective is to compute the kurtosis and frequency of price changes as they appear in a sample generated by this mixture. We focus on an economy in which the size of the measurement error is small, inspired by Eichenbaum et al. (2014), so that we consider the limiting case $\sigma_e \to 0$.[3] Some algebra shows that measured kurtosis is:

$$\lim_{\sigma_e \to 0} Kur(\Delta p_m) \; = \; k \; \frac{\Omega}{\zeta} \qquad \text{where} \qquad \Omega \equiv \frac{\pi_1 \sigma_1^4 + (1-\pi_1)\sigma_2^4}{\left(\pi_1 \sigma_1^2 + (1-\pi_1)\sigma_2^2\right)^2} \; \geq 1 \tag{2}$$

This formula shows that for the estimation to be unbiased, i.e. $Kur(\Delta p_m) = k$, it is necessary that there is no measurement error $\zeta = 1$ and no heterogeneity $\sigma_1 = \sigma_2$. If either of these conditions fails, then measured kurtosis is upward biased. In the empirical analysis we address the heterogeneity-induced bias by standardizing the price changes at a disaggregate cell level (by demeaning the price change observations and dividing them by their standard deviation). A cell is, for instance, a category of good and of outlet type, such as a baguette in

[3]The technical Appendix H.2 explores the concerns about the spurious small price changes raised by Eichenbaum et al. (2014), concluding that they also apply to the French CPI data, albeit to a lesser extent. Even at the finest level of disaggregation some price changes in the CPI data reflect product substitutions (e.g. different brands for a given good being recorded) rather than an actual change in the good's price. Likewise, spurious small price changes may originate in scanner dataset from the unit value problem of weekly prices, whereby the recorded prices average the prices paid by customers with and without discount coupons.

supermarkets. A similar approach was followed by Klenow and Kryvtsov (2008) and Midrigan (2011). Notice that, absent measurement error, the standardization procedure would deliver an unbiased estimate of kurtosis k. To address the small measurement error issue we discard observations with tiny price changes (e.g. smaller than 1 cent), a common practice in the empirical studies of price changes. We then compute the statistics for the pooled standardized data.

The model also illustrates the differential impact of measurement error and heterogeneity on the measured frequency of price changes. Measurement error makes the (expected) number of measured price changes, N_m, greater than the number of true price changes, N_p, so that $N_m = N_p + N_e$ where N_e is the number of incorrectly imputed price changes (all measured per period). This prediction is supported by recent evidence by Cavallo (2015), who shows that price durations estimated over scanner datasets are affected by measurement issues that produce a downward bias of measured durations. We have that $N_e = (1 - \zeta)N_m$ and $N_p = \zeta\, N_m$. Notice that heterogeneity has no effect on the measurement of the frequency of price change, unlike in the case of kurtosis. If two samples are observed, one measurement error free and the other one with a fraction $1 - \zeta$ of small spurious price changes, then ζ can be estimated using the ratio of the two estimated frequencies of price changes. If kurtosis k and k_m is also measured on the two samples, then one can infer what part of the bias in the measurement can be attributed to measurement error (using the bias factor $1/\zeta$) and what part is due to heterogeneity (the factor Ω).

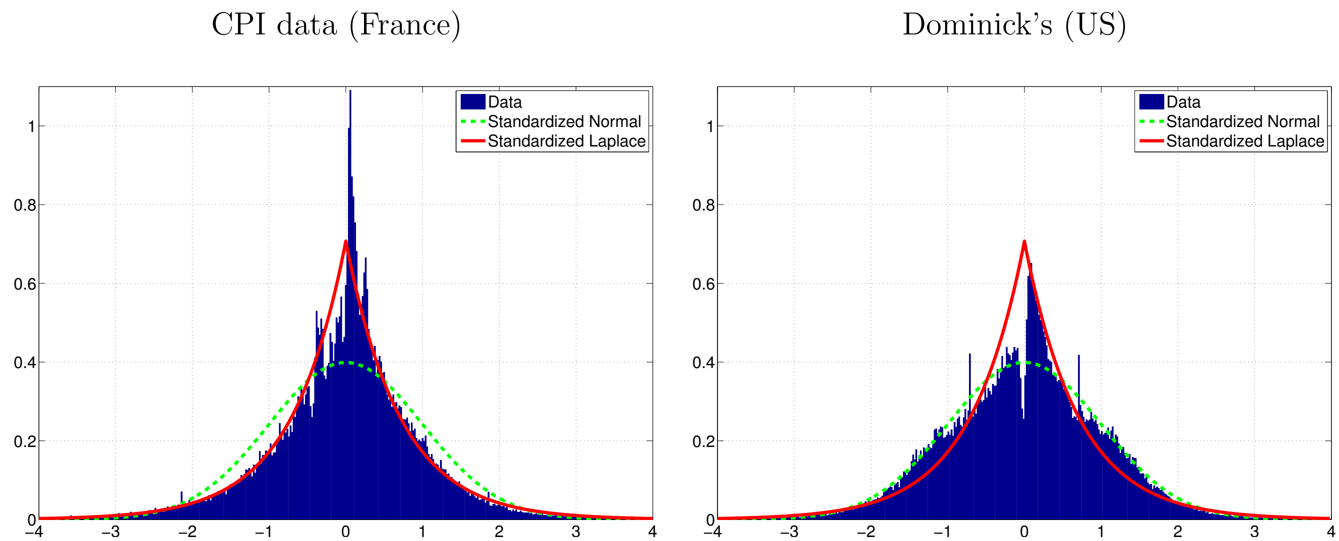
2.2 Evidence

We use two large datasets to provide evidence on the kurtosis of the size distribution, accounting for measurement error and heterogeneity. The first one is the Dominick's dataset, featuring weekly scanner data from a large supermarket chain in Chicago.[4] Around 15,000 UPCs are available, belonging to 29 product categories (such as beer or shampoo), over 400 weeks from September 1989. Following Midrigan we focus on one particular store, the one with most observations (store #122). The second piece of evidence is a longitudinal dataset of monthly price quotes underlying the French CPI, over the period 2003:4 to 2011:4, containing around 11 million price quotes, documented in details in Berardi, Gautier, and Le Bihan (2015). Each record relates to a precisely defined product sold in a particular outlet in a given month. One main advantage of CPI data is the broader coverage of household consumption. The raw dataset covers about 65% of the CPI basket (some categories of goods and services are not available in our sample). The dataset also includes CPI weights, which

[4] Data at the universal product code (UPC) level are provided by James M. Kilts Center, University of Chicago Booth School of Business.

we use to compute aggregate statistics.

Figure 1: Histogram of Standardized Price Changes: France (CPI) and US (Dominick's)



The figures use the elementary CPI data from France (2003-2011), and the Dominick's data set. Price changes are the log difference in price per unit, standardized by good category (272) and outlet type (11) and pooled. Price changes equal to zero are discarded. The panel with French CPI uses about 1.5 million data points, the panel with Dominick's about 0.3 million.

In both datasets price changes are computed as 100 times the log-difference in prices (price per unit in the case of CPI where package size may vary, unlike with UPCs). As a first hedge against measurement error we discard observations with item substitutions in CPI data (which might give rise to spurious price changes). We apply the following trimming to the data: for Dominick's we disregard changes that are smaller than 1 cent (due to measurement error) and drop observations with price levels smaller than 20 cents or larger than 25 dollars (deemed implausible in view of the type of items sold and the distribution of price levels). In the CPI data we drop price changes whose absolute value is smaller than 0.1 percent. In both datasets, we remove as "outliers" observations with log-price changes larger in absolute value than the 99^{th} percentile of absolute log price changes. To handle the issue of heterogeneity, we standardize the data at the "cell" level. A cell is defined by a good category and outlet type in the CPI data, and a UPC-store in scanner data. We use the finest partition possible in our data: for CPI each cell is a COICOP category at the 6-digit level in an outlet type, and we have around 1,500 cells. As mentioned we compute the standardized price changes by subtracting the cell mean (for all non zero price changes) and dividing by the cell-specific standard deviation.

Figure 1 summarizes our main findings with a weighted histogram of the standardized price changes. On the same graph we superimpose the density of the standard Normal distribution as well as the standardized Laplace distribution (both have unit variance). The Laplace distribution has a kurtosis of 6 and is thus more peaked than the Normal. It is apparent that both empirical distributions of standardized price changes are more peaked than the Normal. The kurtosis of the standardized price changes measured on the French CPI is large, equal to 8.0 (removing sales has a minor effect: kurtosis increases to 8.9). This estimate is still likely upward biased because of the remaining measurement errors and the remaining heterogeneity, which increase estimated kurtosis as indicated by $\Omega > 1$ and $1/\zeta > 1$ in equation (2). Notice that our control for heterogeneity in the CPI data is partial, as the information available prevents us from correcting heterogeneity at a finer level: e.g., we do not know the UPC of the product or the store where it is sold. To get a sense of how this residual heterogeneity impacts results, we performed a computation in the Dominick's datasets by standardizing price changes at the product level (i.e. beer, or shampoo which roughly matches the finest level of disaggregation available in the CPI) as opposed to the UPC level. Kurtosis is then 4.9, against 4.0 in the baseline UPC-level standardization. Notice that in addition to this 20% bias, the CPI has the additional problem that CPI prices are not collected in a single store or area, giving rise to yet another source of heterogeneity that is not present in Dominick's data (where we focus on a single store and area). Moreover, measurement error is likely present, as our trimming of the data is quite conservative: using a 1% to identify spurious small price changes as suggested by Eichenbaum et al. (2014) reduces estimated kurtosis to 7 (see the technical Appendix H.1 for more evidence).

As a further assessment of measurement error, we matched a specific subset of the French CPI data with data taken from the Billion Price Project (BPP) dataset, see Cavallo (2015). These data were constructed with the specific intent of addressing heterogeneity and measurement error issues, and so they provide an ideal environment to assess the relevance of the issues discussed above.[5] We matched the BPP data from 3 retailers with the corresponding items in the CPI (see Appendix A). We find that for these retailers kurtosis computed on the BPP data is around 1/2 the value computed from using the CPI data. This suggests that the factor Ω/ζ in equation (2), is around 2 in this sample. Extrapolating this finding to the CPI index, yields an estimated kurtosis of about 4. As discussed in Section 2.1 measurement error (but not heterogeneity) also affects the measurement of the frequency of price changes. Indeed, the data on duration in the BPP vs the CPI is also consistent with the presence of substantial measurement error. In two of the three outlets considered the frequency of

[5] Cavallo (2015) and Cavallo, Neiman, and Rigobon (2014) document that online prices are representative of offline prices for a selected sample of large retailers.

adjustment is half the frequency detected for the corresponding outlet in the CPI data, suggesting that the bias in kurtosis is mostly due to measurement error. In the third outlet by contrast the duration of price adjustment between the BPP vs the CPI is similar, suggesting that the discrepancy in kurtosis mainly reflects residual heterogeneity in the CPI data.

Table 1: Overview of estimates of kurtosis

	US				Other countries		
Source:	M11	NS08	V13	CR15	KR14	W10	CR15
Kurtosis:	4.0	5.1	4.9	4.1	3.98	5.7–8.1	4.0

The kurtosis is computed using standardized price changes. The labels for the various studies are: M11: Midrigan (2011), NS08: Nakamura and Steinsson (2008), V13: Vavra (2014), KR14: Karadi and Reiff (2014) (Hungary) , W10:Wulfsberg (2010) (Norway), CR15: Cavallo and Rigobon (2016) (US and median value across 30 low-inflation countries).

Our analysis of the Dominick's data reveals that kurtosis of the standardized price changes is equal to 4.0. This finding is consistent with the hypothesis that the more granular and more precise nature of the information in the dataset allows for better control over heterogeneity as well as measurement errors. Finally, Table 1 provides an overview of available estimates of the kurtosis of standardized price changes for the US and some other countries. Most estimates are located in the vicinity of 4. In particular, the estimates by Cavallo and Rigobon (2016) based on the BPP internet scraped data discussed above stand out as the most comprehensive (across countries) and least affected by measurement errors, although the coverage of the goods is not as comprehensive as the CPI. These authors use data from 40 countries, standardizing price changes at the UPC level to deal with heterogeneity. Their Table 2 shows that Kurtosis is 4 in the US and in France (values are rounded to integers). The median value in the 40 country sample is 5, while in the subsample of 30 countries with inflation below 5 percent the median is 4.

3 A tractable model with a random menu cost

This section presents a menu cost model aimed at qualitatively matching the patterns documented above. In the canonical menu cost model price adjustments occur when a threshold is hit, so that the implied distribution of price changes fails to generate the small changes that appear in the data (see the discussion in Midrigan (2011); Cavallo (2015); Alvarez and Lippi (2014)). The model that we propose here is able to produce a large mass of small price changes and the positive excess kurtosis that we documented above. Two ingredients are key to this end: (i) random menu costs and (ii) the menu cost faced by the firm, ψ, applies to a bundle of n goods, so that after paying the fixed cost the firm can reprice all goods at no extra cost. Each of these assumptions individually is able to generate some small price changes and higher kurtosis than in a canonical model with $n = 1$ and constant menu costs. The random menu costs is key to generate a positive excess kurtosis in the distribution of price changes. The combination of the two is important: in the models where $n = 1$ (with or without random menu costs) the distribution of price changes has a mass point at the adjustment threshold, a feature that is in stark contrast with the evidence. The prominence of "large" price changes (i.e. a "U shaped" distribution) persists even in a model with $n = 2$, as in Midrigan (2011) where the distribution of price changes asymptotes near the adjustment threshold, or $n = 3$ as in Bhattarai and Schoenle (2014). We show below that in order to generate a size distribution whose shape is comparable to the data one needs $n \geq 6$.

General Equilibrium Setup. The general equilibrium set up is essentially the one in Golosov and Lucas (2007), adapted to multi-product firms (see Appendix B in Alvarez and Lippi (2014) for details). Households have a constant discount rate r and an instantaneous utility function which is additively separable: a CES consumption aggregate, log in real balances, linear in leisure, with constant intertemporal elasticity of substitution $1/\epsilon$ for the consumption aggregate, so that the Marshallian (or uncompensated) labor supply elasticity to real wages is $1/\epsilon - 1$. A convenient implication of this setup is that nominal wages are proportional to the money supply in equilibrium, so that a monetary shock increases the firms' marginal costs proportionately. Each firm produces n goods, each with a linear labor-only technology, subject to idiosyncratic productivity shocks independent across products, whose log follows a brownian motion with instantaneous variance σ^2. The firm faces a demand with constant elasticity $\eta > 1$ for each of its n products, coming from the household's CES utility function for the consumption aggregate. To keep the expenditure shares stationary across goods in the face of the permanent idiosyncratic shocks, we assume offsetting preference shocks (as in Woodford (2009), Bonomo, Carvalho, and Garcia (2010) Midrigan (2011), Alvarez and Lippi (2014)). The frictionless profit-maximizing price for good i at time t is thus given by a constant markup $\eta/(\eta-1)$ over the marginal cost. Let $P_i^*(t)$ be the log of the frictionless profit maximizing price which follows the process $\mathrm{d}P_i^*(t) = \sigma \ \mathrm{d}W(t)$ where $W(t)$ is a standard brownian motion with no drift, and σ is the standard deviation of productivity. The technology to change prices is as follows: each firm is subject to a random menu cost to simultaneously change the price of its n products. In a period of length dt this cost amounts to ψ_L units of labor with probability $1 - \lambda dt$, or zero with probability λdt. Let $p_i(t)$ denote the "price gap" for good i at time t, i.e. the difference between the actual (log) sale price $P_i(t)$ and the (log) profit maximizing price $P_i^*(t)$, i.e. $p_i(t) \equiv P_i(t) - P_i^*(t)$. The firm flow profit can be approximated, up to a second order, by a quadratic loss function $\sum_i^n Bp_i(t)^2$ where the scale parameter $B = (1/2)\eta(\eta - 1)$ is related to the demand elasticity.[6] Under this approximation it is convenient to express the menu cost ψ_L in units of flow profits at the optimal price, a constant value that we denote by ψ.

To keep the model simple we assume that all goods in the economy are ex-ante symmetric and subject to shocks with a common variance σ^2, and all firms share the same fundamental parameters. An extension to a case with ex-ante heterogeneous firms, involving different frequencies and kurtosis of price adjustment, is presented in Appendix E. Our empirical analysis of heterogeneity and the standardization of data employed in the previous section are fully consistent with the theoretical framework discussed in this extension provided that the shocks that hit the individual goods are normally distributed (but possibly with different variances).

3.1 A simple case with $n = 1$ good.

The simplest illustration of our random menu cost model obtains for the case where $n = 1$, so the price gap p is scalar. Let $V(p)$ be the present-value cost function for the firm. Upon the arrival of a free adjustment opportunity, i.e. a zero menu cost, the firm optimally resets the price gap to zero (given the symmetry of the loss function and the law of motion of price gaps), hence the Bellman equation for the range of inaction reads:

$$r\ V(p) = Bp^2 + \lambda\Big(V(0) - V(p)\Big) + \frac{\sigma^2}{2}\ V''(p)\ , \quad \text{for } p \in (-\bar{p}, \bar{p})\ ,$$

where $\bar{p}$ is the threshold rule defining the region where inaction is optimal (see the technical Appendix I for the calculations of this section). This equation states that the flow value of the Bellman equation is given by the instantaneous losses, Bp^2, plus the expected change in the value function, which is due either to a free adjustment (with rate λ in which case the price gap is reset to zero) or to the volatility of shocks σ^2. The value-matching and smooth-pasting conditions are given by $V(\bar{p}) = V(0) + \psi$ and $V'(\bar{p}) = 0$. A Taylor expansion of the value function yields the following approximate optimal threshold $\bar{p} = \left(\frac{6\psi\sigma^2}{B}\right)^{\frac{1}{4}}$ which is accurate for small values of the menu cost ψ.[7]

Computing the expected time between adjustments yields an expression for the average

[6]This approximation is quite accurate for the typical small values of the menu cost used in the literature. See Alvarez and Lippi (2014) for a quantitative illustration of the accuracy and for a survey of several papers that use the quadratic formulation.

[7] The approximation obtains as $\frac{\psi(r+\lambda)^2}{B\sigma^2} \to 0$, see Proposition 3 of Alvarez and Lippi (2014). Exactly the same expression was established by Barro (1972) and Dixit (1991) for the case in which $\lambda = 0$. Below we discuss an approximate threshold for large values of ψ that is useful to interpret the Calvo model.

number of adjustments per period, $N(\Delta p_i)$, which we use to measure the fraction of free adjustments over the total number of adjustments, a variable we call ℓ, as

$$\ell \equiv \frac{\lambda}{N(\Delta p_i)} = \frac{e^{\sqrt{2\phi}} + e^{-\sqrt{2\phi}} - 2}{e^{\sqrt{2\phi}} + e^{-\sqrt{2\phi}}} \in (0,1) \qquad \text{where we define} \qquad \phi \equiv \frac{\lambda \bar{p}^2}{\sigma^2}$$

which shows that the fraction of free adjustments ℓ depends only on the parameter ϕ. The parameter ϕ can be interpreted as the ratio between λ, the number of free adjustments, and $\sigma^2/\bar{p}^2$, the number of adjustments in a model where $\lambda = 0$ and the threshold policy $\bar{p}$ is followed.

Let $\Delta p_i = -p$ denote the price change implemented by a firm that adjusts when its price gap is p. The distribution of price changes is symmetric around $\Delta p_i = 0$. This distribution has two mass points at $\Delta p_i = \pm\bar{p}$. The two points, which account for $1 - \ell$ of the mass, are due to the price changes that occur when the price gap hits the boundaries of the inaction region. The remaining price changes, a fraction ℓ of the mass, occur when a free adjustment opportunity arrives, at which time the price gap is set to zero. Price changes in the range $\Delta p_i \in (-\bar{p}, \bar{p})$ have a density $\ell\, g(p)$ where $g(p)$ denotes the density of the invariant distribution of price gaps given by

$$g(p) = \frac{\sqrt{2\phi}}{2\bar{p}\left(e^{\sqrt{2\phi}} - 1\right)^2} \left(e^{\sqrt{2\phi}\left(2 - \frac{|p|}{\bar{p}}\right)} - e^{\sqrt{2\phi}\frac{|p|}{\bar{p}}}\right) \quad \text{for} \quad p \in [-\bar{p}, \bar{p}] \,. \tag{3}$$

This density is a symmetric in the $(-\bar{p}, \bar{p})$ interval and its shape is high-peaked ("tent-shaped"). The full distribution features two mass points at the boundaries $\pm\bar{p}$. As shown below the kurtosis of this distribution is increasing in λ, and in particular the distribution of price changes is more peaked than that of the standard menu cost model where $\lambda = 0$.

We notice that the shape of the distribution of price changes depends *only* on the fraction of free adjustments ℓ (or, equivalently, on ϕ). This means that two economies, or sectors, that differ in the standard deviation of price changes $Std(\Delta p_i)$ and/or in the frequency of price adjustment $N(\Delta p_i)$ will display a distribution of price changes with exactly the same shape (once its scale is adjusted) provided that they have the same value of ℓ. This property is useful to aggregate the sectors of an economy that are heterogenous in their steady state features $N(\Delta p_i)$, $Std(\Delta p_i)$. Because kurtosis is a scale free statistic, i.e. independent of $Std(\Delta p_i)$, it is completely determined by ϕ in this model. The general model adds one parameter, n, as a determinant of the shape of the size distribution of price changes and hence of kurtosis.

3.2 The model with multi-product firms

This section incorporates the model with free adjustment opportunities discussed above into the model of Alvarez and Lippi (2014) where the firm is selling n goods, so that p is now a vector in $\mathbb{R}^n$, but pays a single fixed adjustment cost to change the n prices. We incorporate this feature for several reasons. First, as explained above, in the model with $n = 1$ good there is a mass point on price changes of size $|\Delta p_i| = \bar{p}$. For $n = 2$ the mass point disappears but the distribution of price changes still features the largest mass of observations (the highest density) near the adjustment thresholds. There is no evidence of this in any data set we can find. Values of $n \geq 6$ produce a bell-shaped size distribution of price changes that is much closer to what is seen in the data. Second, the model with $\lambda = 0$ has a kurtosis that increases with n, hence providing an alternative to random menu costs. Third, for large n and $\lambda = 0$ the distribution of price changes tends to the Normal distribution, which is both a nice benchmark and an accurate description of the price changes for some sectors. Finally, the multi-product model with $(n > 1)$ has an alternative, broader, "rational inattentiveness" interpretation for the adjustment cost ψ. In particular, one can assume that the firm freely observes its total profits but not the individual ones (for each product), unless it either pays the cost ψ or a free observation opportunity arrives, in which case it is able to set the optimal price to each of them. This allows a broader interpretation of the menu cost, including not only the physical cost of changing prices but also the cost related to gathering and processing the information for individual products.[8] For instance, as $n \to \infty$ the model converges to the rational inattentiveness model of Section 4.3 in Reis (2006).

We now briefly describe the setup of the firm problem with n products. As before the free adjustment opportunities are independent of the driving processes $\{W_i(t)\}$ for price gaps $i = 1, 2, ..., n$, and arrive according to a Poisson process with constant intensity λ. In between price adjustments each of the price gaps evolves according to a Brownian motion $\mathrm{d}p_i(t) = \sigma \, \mathrm{d}W_i(t)$. It is assumed that all price gaps are subject to the same variance σ^2 and that the innovations are independent across price gaps.[9]

We assume that, when the free opportunity arrives, the firm can adjust all prices without paying the cost ψ. The analysis of the multi product problem can be greatly simplified by using the sum of the squared price gaps, $y \equiv ||p||^2$ as a state variable, instead of the vector $p = (p_1, ..., p_n)$, as done in Alvarez and Lippi (2014). The scalar y summarizes the state

[8] As an example, see Chakrabarti and Scholnick (2007) who argue that for stores such as Amazon or Barnes and Noble physical menu cost are small, yet prices change infrequently, and thus conclude that the cost may be of a different nature. Interestingly, they find that for such retailers price changes are synchronized across products, which is an implication of the multi-product model.

[9] Alvarez and Lippi (2014) discuss the case with correlated price gaps. Intuitively, as correlation increases the model becomes more similar to the $n = 1$ case, since the price gaps of a firm become more similar.

because the period objective function can be written as a function of it and because, from an application of Ito's lemma, one can derive a one dimensional diffusion which describes its law of motion, namely

$$\mathrm{d}y = n\sigma^2 \,\mathrm{d}t \,+\, 2\,\sigma\sqrt{y}\,\mathrm{d}W$$

where W is a standard Brownian motion.

Using $N(\Delta p_i)$ and $Var(\Delta p_i)$ to denote the frequency and the (cross sectional) variance of the price changes of product i, the next proposition establishes a useful relationship that holds in a large class of models for *any policy* for price changes, which we describe by a stopping time rule:

Proposition 1 *Let τ describe the time at which a price change takes place, so that all price gaps are closed. Assume the stopping time treats each of the n price gaps symmetrically. For any finite stopping time τ we have:*

$$N(\Delta p_i) \cdot Var(\Delta p_i) = \sigma^2 \ . \tag{4}$$

The proposition highlights the trade-off for the firm's policy: more frequent adjustments are required to have smaller price gaps. See Appendix B for the proof, where the reader can verify that the key assumptions are random walks and symmetry. We underline that equation (4) holds for any stopping rule, not just for the optimal one. Indeed it holds for a larger class of models, for instance those with correlated price gaps and a richer class of random adjustment cost.

Upon the arrival of a free adjustment opportunity the firm will set the price gap to zero, hence the Bellman equation for the range of inaction reads:

$$r\ v(y) = B\,y + \lambda\Big(v(0) - v(y)\Big) + n\sigma^2\ v'(y) + 2\sigma^2 y\ v''(y), \qquad \text{for } y \in (0, \bar{y})\ , \tag{5}$$

where $B\,y$ is the sum of the deviation from the optimal profits from the n goods.

Given the symmetry of the problem after an adjustment of the n prices the firm will set all price gaps to zero, i.e. will set $||p||^2 = y = 0$. The value matching condition is then $v(0) + \psi = v(\bar{y})$, which uses the fact that when y reaches a critical value, denoted by $\bar{y}$, the firm can change the n prices by paying the fixed cost ψ. The smooth pasting condition is $v'(\bar{y}) = 0$.

The next lemma establishes how to solve for $\bar{y}$ using the solution of the problem with $\lambda = 0$ discussed in Alvarez and Lippi (2014). In particular, using $r + \lambda$ as a modified interest rate in the solution of the problem with $\lambda = 0$, allows us to immediately compute the solution for the case of interest in this paper. We have:

Lemma 1 *Let $v(y; r, \lambda)$ and $\bar{y}(r, \lambda)$ be the optimal value function and adjustment threshold for a problem with discount rate r and arrival rate λ. Then $v(y; r, \lambda) = v(y; r + \lambda, 0) + \frac{\lambda}{r} v(0; r + \lambda, 0)$ for all $y \geq 0$ and thus $\bar{y}(r, \lambda) = \bar{y}(r + \lambda, 0)$.*

The proof of this lemma follows a straightforward guess and verify strategy. The lemma allows us to use the characterization of $\bar{y}$ with respect to r given in Proposition 4 of Alvarez and Lippi (2014) to study the effect of $r + \lambda$ on $\bar{y}$.[10] The next proposition summarizes that result and extends the characterization of the optimal threshold to the case where ψ is large, a case that is useful to understand the behavior of an economy with a lot of free adjustments opportunity as in a Calvo mechanism (see Appendix B for the proof).

Proposition 2 *Assume $\sigma^2 > 0$, $n \geq 1$, $\lambda + r > 0$ and $B > 0$, and let $\bar{y}$ be the threshold for the optimal decision rule. We then have that:*

1. *As $\psi \to 0$ then $\frac{\bar{y}}{\sqrt{2(n+2)\sigma^2 \frac{\psi}{B}}} \to 1$ or $\bar{y} \approx \sqrt{2(n+2)\sigma^2 \frac{\psi}{B}}$.*
2. *As $\psi \to \infty$ we have $\frac{\bar{y}}{\psi} \to (r + \lambda)/B$ or $\bar{y} \approx \frac{\psi}{B}(r + \lambda)$. Moreover this also holds for large n and large $\frac{\psi}{n}$, namely $\lim_{\psi/n \to \infty} \lim_{n \to \infty} \frac{\bar{y}/n}{\psi/n} = (r + \lambda)/B$ or $\frac{\bar{y}}{n} \approx \frac{\psi/n}{B}(r + \lambda)$.*

The proposition shows that $\bar{y}$ is approximately constant with respect to λ for small values of ψ, so that for small menu costs the result is the well known quartic root formula (recall that y has the units of a squared price gap) and the inaction region is increasing in the variance of the shock, due to the higher option value. Interestingly, and novel in the literature, the second part of the proposition shows that for large values of the adjustment cost the rule becomes a square root and that the optimal threshold does not depends on σ, which shows that for large adjustment costs the option value component of the decision becomes negligible. Moreover, when the menu costs are large the threshold $\bar{y}$ is increasing in λ: the prospect of receiving a free adjustment tomorrow increases inaction today.

We now turn to the discussion of the model implications for the frequency of price changes. We let $N(\Delta p_i)$ be the expected number of adjustments per unit of time of a model with a given λ and $\bar{y}$. We establish the following (see Appendix B for the proof):

Proposition 3 *Let Γ denote the gamma function. The fraction of free adjustments is $\ell = \lambda / N(\Delta p_i)$, where*

$$\ell = \frac{\sum_{i=1}^{\infty} \frac{\Gamma\left(\frac{n}{2}\right)}{i!\,\Gamma\left(\frac{n}{2}+i\right)} \left[\frac{n}{2}\right]^i \phi^i}{\sum_{i=0}^{\infty} \frac{\Gamma\left(\frac{n}{2}\right)}{i!\,\Gamma\left(\frac{n}{2}+i\right)} \left[\frac{n}{2}\right]^i \phi^i} \equiv \mathcal{L}(\phi, n), \quad \textit{where } \phi \equiv \frac{\lambda \bar{y}}{n\sigma^2} \tag{6}$$

[10] The technical Appendix K gives the analytical solution for the value function and provides more details.

The proposition shows that ℓ is a function only of two variables: n and ϕ, and that it is increasing in ϕ. As was the case for the model with $n = 1$, the parameter ϕ can be interpreted as the ratio between λ, the number of free adjustments, and $n\sigma^2/\bar{p}^2$, the number of adjustments in a model where $\lambda = 0$ and the threshold policy $\bar{y}$ is followed. For a given n there is a one to one and onto mapping between ϕ and ℓ: as $\phi \to 0$ then $\ell \to 0$, and as $\phi \to \infty$ then $\ell \to 1$.

Finally we characterize the invariant distribution of y for the case where $\lambda > 0$, a key ingredient to compute the size-distribution of price changes. The density of the invariant distribution solves the Kolmogorov forward equation: $\frac{\lambda}{2\sigma^2} f(y) = f''(y)y - \left(\frac{n}{2} - 2\right) f'(y)$ for $y \in (0, \bar{y})$, with the two boundary conditions $f(\bar{y}) = 0$ and $\int_0^{\bar{y}} f(y)dy = 1$. It is clear from these conditions that $f(\cdot)$ is uniquely defined for a given triplet: $\bar{y} > 0$, $n \geq 1$ and $\lambda/\sigma^2 \geq 0$. The general solution of this ODE is

$$f(y) = \left(\frac{\lambda y}{2\sigma^2}\right)^{\left(\frac{n}{4}-\frac{1}{2}\right)} \left(C_1 \ I_\nu\left(2\sqrt{\frac{\lambda y}{2\sigma^2}}\right) + C_2 \ K_\nu\left(2\sqrt{\frac{\lambda y}{2\sigma^2}}\right)\right) \tag{7}$$

where I_ν and K_ν are the modified Bessel functions of the first and second kind, C_1, C_2 are two arbitrary constants and $\nu = |\frac{n}{2} - 1|$, see Zaitsev and Polyanin (2003) for a proof. The constants C_1, C_2 are chosen to satisfy the two boundary conditions.[11] While the density in equation (7) depends on 3 constants n, ϕ and $\bar{y}$, its shape depends only on 2 constants, namely n and ϕ, as formally stated in Lemma 3 in Appendix B. The lemma shows that one can normalize $\bar{y}$ to 1 and compute the density for the corresponding ϕ.

We denote the marginal distribution of price changes by $w(\Delta p_i)$. Recall that a firm changes all prices when y first reaches $\bar{y}$ or when a free adjustment opportunity occurs even though $y < \bar{y}$. Therefore to characterize the price changes Δp_i of good i belonging to the vector of price gaps p we need three objects: the fraction of free adjustments ℓ, the invariant distribution $f(y)$ and the marginal distribution of price changes *conditional on a value of* y, $\omega(\Delta p_i; y)$ which, following Proposition 6 of Alvarez and Lippi (2014) when $n \geq 2$, is

$$\omega(\Delta p_i; y) = \begin{cases} \frac{1}{Beta\left(\frac{n-1}{2}, \frac{1}{2}\right) \sqrt{y}} \left(1 - \left(\frac{\Delta p_i}{\sqrt{y}}\right)^2\right)^{(n-3)/2} & \text{if} \quad (\Delta p_i)^2 \leq y \\ 0 & \text{if} \quad (\Delta p_i)^2 > y \end{cases} \tag{8}$$

where $Beta(\cdot, \cdot)$ denotes the Beta function. In this case the (cross-sectional) standard deviation of the price changes is $Std\,(\Delta p_i\ ; y) = \sqrt{y/n}$. The marginal distribution of price changes

[11] We note that both modified Bessel functions are positive, that $I_\nu(y)$ is exponentially increasing with $I_\nu(0) \geq 0$, and that $K_\nu(y)$ is exponentially decreasing with $K_\nu(0) = +\infty$. The technical Appendix O gives a closed form expression for f in terms of power series.

$w(\Delta p_i)$ is given by

$$w(\Delta p_i) = \omega(\Delta p_i; \bar{y})\,(1-\ell) + \Big(\int_0^{\bar{y}} \omega(\Delta p_i; y) f(y) dy\Big)\ell \qquad \text{for } n \geq 2 \quad . \tag{9}$$

which shows that the distribution $w(\Delta p_i)$ is a mixture of the $\omega(\Delta p_i, y)$ densities. These densities are scaled versions of each other with different standard deviations. The mixture increases the kurtosis of the distribution of price changes compared to the case where $\lambda = 0$, for the reasons discussed in Section 2.[12] Before illustrating the shapes produced by this distribution, the next proposition shows that those shapes are completely determined by n and ℓ and no other parameters (see Appendix B for the proof):

Proposition 4 *Let $w(\Delta p_i; n, \ell, 1)$ be the density function for the price changes Δp_i in an economy with n goods, a share ℓ of free adjustments, and a unit standard deviation of price changes $Std(\Delta p_i) = 1$. This density function is homogenous of degree -1 in Δp_i and $Std(\Delta p_i)$, which implies*

$$w\,(a\,\Delta p_i; n, \ell, a) = \frac{1}{a} \;\; w\,(\Delta p_i; n, \ell, 1) \qquad \textit{for all} \quad a > 0. \tag{10}$$

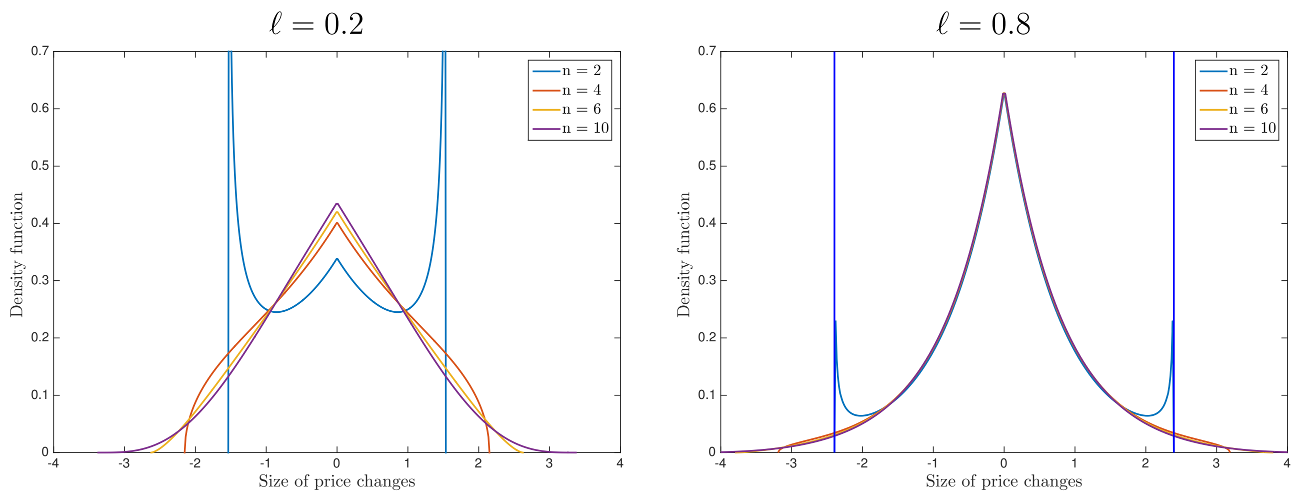
The proposition implies that we can aggregate firms or industries that are heterogenous in terms of frequency $N(\Delta p_i)$ and standard deviation of price changes $Std(\Delta p_i)$ provided that n and ℓ are the same. Notice in particular that the frequency of price changes $N(\Delta p_i)$ does not have an independent effect on the distribution of price changes as long as ℓ remains constant.

Figure 2 shows the shapes of the distribution of price changes Δp_i in equation (9) obtained for different combinations of n and ℓ. It is evident that for small values of n the shape of the distribution does not match the "tent-shaped" patterns that are seen in the data as in e.g. Figure 1. For the case when $n = 2$ the density of the price changes diverges at the boundaries of the domain where $\Delta p_i = \pm\sqrt{\bar{y}/n}$. This feature echoes the two mass points that occur in the $n = 1$ case where a non-zero mass of price changes occurs exactly at the boundaries. For $n \geq 6$ the shape of the density takes a tent-shape, similar to the one that is seen in the data. As the fraction of free adjustments approaches 1 the density function converges to the Laplace distribution.

Using that Δp_i is distributed as a mixture of the $\omega(\Delta p_i, y)$, we can compute several

[12] In particular Proposition 6 in Alvarez and Lippi (2014) shows that the variance and kurtosis of $\omega(\Delta p_i, y)$ are given by y/n and $3n/(n + 2)$ respectively.

Figure 2: Size distribution of price changes



Note: All distributions are zero mean with unit standard deviation. As stated in Proposition 4 the shape of this distribution only depends on ℓ and n.

moments of interest, such as

$$\begin{aligned} Var(\Delta p_i) &= (1-\ell)\,\frac{\bar{y}}{n} + \ell \int_0^{\bar{y}} \frac{y}{n}\; f(y)dy \qquad (11) \\ Kur(\Delta p_i) &= \frac{3n}{2+n}\;\frac{(1-\ell)\,\bar{y}^2 + \ell\;\int_0^{\bar{y}} y^2\; f(y)dy}{\Big((1-\ell)\,\bar{y} + \ell\;\int_0^{\bar{y}} y\; f(y)dy\Big)^2} > \frac{3n}{2+n} \end{aligned}$$

As stated in Proposition 4 the value of the kurtosis $Kur(\Delta p_i)$ depends only on two parameters: n and ℓ. Moreover, kurtosis is increasing in both n and ℓ, as can be seen in Figure 4 which plots the value of $Kur(\Delta p_i)/6$ for various combinations of n and ℓ, the only two parameters determining kurtosis. For small values of ℓ kurtosis is increasing in n up to a level of 3. For instance, if $\ell = 0$ and $n \to \infty$, the kurtosis converges to 3 since the distribution of price changes at the time of adjusting for each firm becomes normal; this value is the highest that the *purely* multi-product model with $\ell = 0$ can produce. For any n, as the fraction of free adjustments ℓ increases, the kurtosis increases towards 6, the maximum value achieved in our model when $\ell = 1$.[13] The inequality that appears in the second line is a well known result: the mixture of distributions with the same kurtosis but with different

[13]The relationship between the fraction of free adjustment and the "mass of small price changes", a statistic that is closely related to kurtosis, was noticed in the numerical simulations of Nakamura and Steinsson (2010) (see their footnote 15, where their "frequency of low repricing opportunities", $1-\alpha$, is essentially our ℓ).

variances has higher kurtosis, which follows from Jensen's inequality.[14]

To conclude the description of the model we summarize a few special cases nested by our setup. The Golosov-Lucas model is obtained when $n = 1$ and $\ell = 0$, implying a kurtosis of 1. The Taylor model, or equivalently "rational inattentiveness" model by Reis, is obtained when $n = \infty$ and $\ell = 0$, implying a kurtosis of 3. The Calvo model is obtained for $\ell \to 1$, for all values of n, implying a kurtosis of 6. Additionally, two recent models can be proxied: the "CalvoPlus" model by Nakamura and Steinsson (2010) for the special case of no intermediate goods ($s_m = 0$ using their notation), is obtained assuming $n = 1$ and $\ell \in (0, 1)$. The multiproduct model of Midrigan (2011) is obtained assuming $n = 2$ and modeling the fat tailed shock by assuming $\ell \in (0, 1)$. Depending on the parameter choice for ℓ, the last two models can generate a kurtosis between 1 and 6. It is shown in Section 4.1 that higher values of kurtosis are essential in both models to explain why the real effects in those models are closer to Calvo than to Golosov-Lucas.

The map between the fundamental parameters and observables. Our model has four independent parameters: the scaled menu cost ψ/B, the volatility of shocks σ, the number of goods n and the rate of free adjustment opportunities λ. The model provides an invertible mapping between these parameters and 4 observables, which is useful to think about parameters' identification and the model's comparative statics. Since different papers, e.g. Golosov-Lucas and Midrigan, have in common that they are matched to the same frequency and standard deviation of price changes for the US, we find it convenient to illustrate the behavior of our model while keeping the frequency $N(\Delta p_i)$ and the variance $Var(\Delta p_i)$ constant. Matching these statistics implies that 2 of the 4 fundamental model parameters, namely ψ/B and σ, are pinned down. The model has two residual parameters: n and λ, with the latter mapping one-to-one and onto $\ell = \lambda/N(\Delta p_i)$. The parametrization of the model can thus be usefully interpreted as choosing n and ℓ to match two additional empirical moments. It was shown in Proposition 4 how ℓ and n shape the distribution of price changes, in particular its kurtosis. In Appendix C we show how ℓ and n map into the cost of price adjustments, for given values of $N(\Delta p_i)$ and $Var(\Delta p_i)$. Kurtosis and the cost of price adjustments can thus be used to discipline the parameterization of the model.

4 The real output effect of a monetary shock

In this section we discuss the response of the economy's aggregate output to an unexpected (once and for all) increase of the money supply of size δ, starting from a steady state with

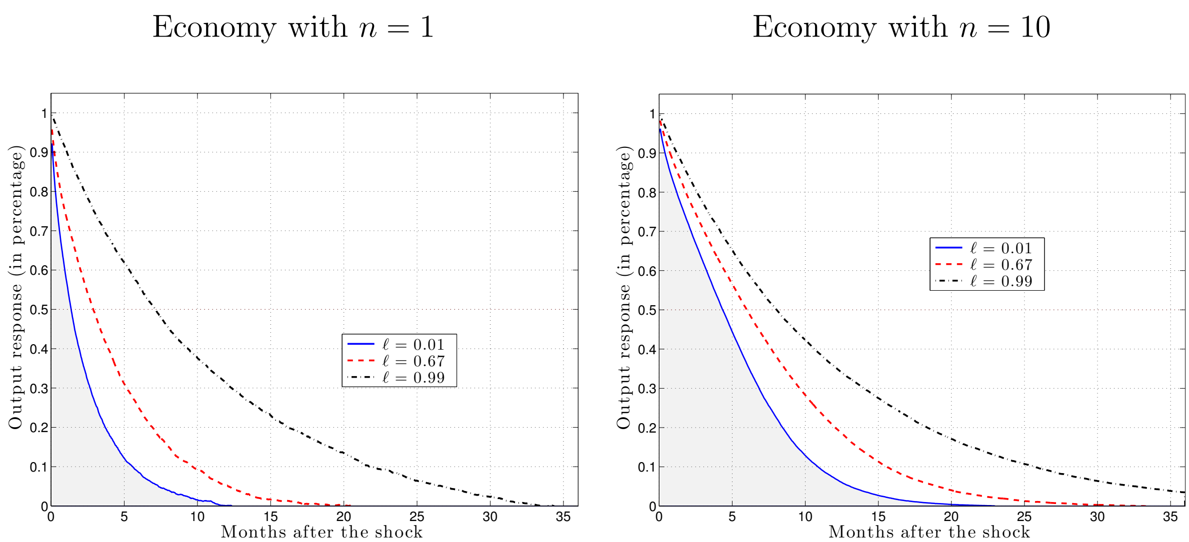
[14] The technical Appendix L gives a closed form solution for kurtosis without solving for $f(y)$.

zero, or small, inflation. Figure 3 plots the impulse response function (IRF) of output produced by our model for different combinations of n and ℓ. The impulse response functions were computed numerically using the decision rules described above. Appendix D gives the derivation in details and establishes analytically two useful results: first that the shape of the IRF depends *only* on n and ℓ. Second, changes in the standard deviation of price changes $std(\Delta p_i)$ and in the frequency of price changes $N(\Delta p_i)$ imply a rescaling of the vertical and horizontal axes, respectively, but do not affect the shape of the IRF. For instance, while the figure is drawn for an economy with $N(\Delta p_i) = 1$, it can be readily used to analyze an economy with $N(\Delta p_i) = 2$ simply by dividing values on the time axis by 2. The left panel of the figure presents the IRF for an economy with $n = 1$. Three cases are given: the $\ell = 0.01$ is essentially the economy of Golosov and Lucas (2007), where the real effects of monetary shocks are short-lived. The outer line corresponding to $\ell = 0.99$ is the Calvo model, where (almost) all adjustments are triggered by exogenous arrivals of adjustment opportunities. The right panel of the figure plots an economy with $n = 10$. Compared to the economy with $n = 1$ the multi-product economy implies a size distribution of the price changes that features several small and large price changes (and no mass points), an arguably realistic feature of the model. It is also apparent that the economy with $n = 10$ implies more persistent effects of monetary shocks, for instance the half-life of the shock roughly doubles that in Golosov and Lucas. The line corresponding to $\ell = 2/3$ is calibrated to match a kurtosis of the size distribution of price changes equal to 4, an empirically plausible value as discussed in Section 2.2. Notice that this calibration, which we see as realistic, delivers a significant increase of the persistence of monetary shocks compared. For instance in the Golosov-Lucas model ($n = 1, \ell = 0$, left panel) the half-life of the shock is close to 1.5 months, while in our calibrated model ($n = 10, \ell = 0.67$, right panel) the half-life is about 6 months. The calibrated model thus produces real effects that are much more persistent than the canonical model, but less persistent than a Calvo model ($\ell = 1.0$) where the half-life is around 8 months.

Our main objective in this section is to characterize the real output effect of monetary shocks using a simple summary statistic, namely the total cumulative output. This statistic measures the area under the output's impulse response function, e.g. the gray shaded area that appears for illustrative purposes in Figure 3 for the models with $\ell = 0.01$. We find this statistic convenient for two reasons. First, it combines in a single value the persistence and the size of the output response, and it is closely related to the output variance due to monetary shocks, which is sometimes used in the literature.[15] Second for small monetary shocks (like the ones typically considered in the literature) this statistic is completely encoded

[15] For more discussion and evidence on the equivalence between the area under the impulse response function and the variance due to monetary shocks see footnote 21 of Nakamura and Steinsson (2010).

Figure 3: Output response to a monetary shock of size $\delta = 1\%$



The figures represent an economy with $N(\Delta p_i) = 1.0$ and $std(\Delta p_i) = 0.10$.

in the frequency of price changes $N(\Delta p_i)$ and the kurtosis of price changes $Kur(\Delta p_i)$, as highlighted by equation (1). These two sufficient statistics thus provide a straightforward metric to compare the workings of different models.

Formally, the cumulative output $\mathcal{M}$ after a shock δ is given by:

$$\mathcal{M}(\delta) = \frac{1}{\epsilon} \int_0^\infty (\delta - \mathcal{P}(\delta, t)) \, dt \tag{12}$$

where $\mathcal{P}(\delta, t)$ is the aggregate price level t periods after the shock. The argument of the integral gives the aggregate real wages at time t, which are then mapped into output through $1/\epsilon$, a parameter related to the (uncompensated) labor supply elasticity. Integrating over time gives the total cumulative real output.

To characterize $\mathcal{M}(\delta)$ we consider the expected cumulative output deviation from steady state of a firm with a vector of price gaps p:

$$m(p_1, ..., p_n) = -\mathbb{E}\left[\int_0^\tau \sum_{i=1}^n p_i(t) \, dt \;\middle|\; p(0) = p\right] \tag{13}$$

where τ is the stopping time associated with the optimal decision rule described as the first time that $||p(t)||^2$ reaches threshold $\bar{y}$ or that a free adjustment opportunity arrives. Note that, by definition, if $||p||^2 \geq \bar{y}$ then $m(p) = 0$. Intuitively, a firm with a price gap $-p_i(t)$ for good type i is producing $p_i(t)/\epsilon$ percent excess output compared to its steady state at time

t. Thus, integrating over time until the (random) time of adjustment τ, $m(p)/n$ gives the expected total excess output produced by a firm that has a vector of price gaps equal to p immediately after the monetary shock. Three remarks about this solution are in order. First, given the GE structure, identical to the one in Golosov and Lucas (2007), a once and for all increase in money supply is followed by a once and for all increase in nominal wages, and leaves nominal interest rates unaltered. Second, by using the steady state decision rule, $\bar{y}$, we are ignoring the general equilibrium feedback effects. In Proposition 7 of Alvarez and Lippi (2014) we showed that, given a combination of the general equilibrium set-up in Golosov and Lucas (2007) and the lack of the strategic complementarities, these effects are of second order.[16] Third, we use the fact that *after* the first price change the expected contribution to output of each firm is zero since positive and negative output contributions are equally likely, i.e. $m(0) = 0$. This is convenient since it allows us to characterize the propagation of the monetary shocks without having to keep track of the time evolution for the whole distribution of price gaps.

The function $m(p)$ defined in equation (13) is extremely useful: exploiting the law of motion of the state yields a differential equation that fully characterizes the cumulative output without having to solve (or simulate) the whole impulse response function. The idea follows the same logic used to compute expected values using a Bellman (or Kolmogorov) equation. For example in the $n = 1$ case, where p is a scalar price gap in $(-\bar{p}, \bar{p})$, a Bellman equation type of logic gives $\lambda m(p) = -p + m''(p)\sigma^2/2$ with boundary condition $m(\bar{p}) = 0$ and negative symmetry $m(p) = -m(-p)$, with closed form solution:

$$m(p) = -\frac{p}{\lambda} + \frac{\bar{p}}{\lambda}\left(\frac{e^{\sqrt{2\phi}\frac{p}{\bar{p}}} - e^{-\sqrt{2\phi}\frac{p}{\bar{p}}}}{e^{\sqrt{2\phi}} - e^{-\sqrt{2\phi}}}\right) \quad \text{for all} \;\; p \in [-\bar{p}\,,\, \bar{p}] \qquad \text{where} \quad \phi \equiv \lambda\bar{p}^2/\sigma^2. \tag{14}$$

The final element to define $\mathcal{M}(\delta)$ is the density of the invariant distribution $g(p)$ for a vector of price gaps $p \in \mathbb{R}^n$ which is directly implied by the invariant density of the squared price gaps $f(y)$, given in equation (7), and by the observation that in steady state the distribution of price gaps with $||p||^2 = y$ is uniform on the n dimensional hypersphere of square radius y, whose closed form expression is given by:

$$g\left(p_1, ..., p_n\right) = f\left(p_1^2 + \cdots + p_n^2\right) \frac{\Gamma\left(n/2\right)}{\pi^{n/2}\;\left(p_1^2 + \cdots + p_n^2\right)^{(n-2)/2}} \quad . \tag{15}$$

[16] This proposition can be extended in a straightforward way to this paper using the logic of Lemma 1.

Thus we can write

$$\mathcal{M}(\delta) = \frac{1}{\epsilon\, n} \int \ldots \int m\left(p_1 - \delta, \ldots, p_n - \delta\right)\, g\left(p_1, \ldots, p_n\right)\, dp_1 \ldots dp_n \tag{16}$$

Note that $\mathcal{M}$ takes $g(p)$ firms with price gap vector p in steady state, shifts them down by δ, which amounts to increasing the marginal nominal cost on impact, and then computes their contribution $m\left(p_1 - \delta, \ldots, p_n - \delta\right)$.

Recall that, for any $n \geq 1$ we can index the steady state of an economy by three numbers, $N\left(\Delta p_i\right)$, $Std\left(\Delta p_i\right)$, and ℓ, for which we can always find the values of $(\lambda, \psi, \sigma^2)$ to rationalize them. The next proposition shows that $\mathcal{M}$ can be normalized, i.e. written in terms of the values of an economy with one price change per year, and where the monetary shock is measured in terms of steady-state size of price changes (see Appendix B for the proof)

Proposition 5 *Consider an economy whose firms produce $n > 1$ products, with steady-state statistics $(N(\Delta p_i), Std(\Delta p_i), \ell)$. The following rescaling of $\mathcal{M}$ holds:*

$$\mathcal{M}\left(\delta; N(\Delta p_i), Std(\Delta p_i), n, \ell\right) = \frac{Std(\Delta p_i)}{N(\Delta p_i)}\, \mathcal{M}\left(\frac{\delta}{Std(\Delta p_i)}; 1, 1, n, \ell\right) \tag{17}$$

Equation (17) shows a useful homogeneity property of $\mathcal{M}$: keeping (n, ℓ) fixed, $\mathcal{M}$ can be scaled by the steady state frequency of price changes $N\left(\Delta p_i\right)$, and that the size of the monetary shock measured relative to steady-state size of price changes $Std\left(\Delta p_i\right)$. This is convenient since, given the steady state features of an economy, $\left(Std\left(\Delta p_i\right), N\left(\Delta p_i\right)\right)$, there are only 2 remaining parameters to characterize the cumulative impulse response: n and ℓ.

4.1 The case of a small monetary shock

To focus on a small shock δ, a realistic standard in this literature, we take the first order approximation to equation (12). Using equation (17) we obtain $\mathcal{M}\left(\delta; N(\Delta p_i), Std(\Delta p_i), n, \ell\right) \approx \delta / N(\Delta p_i)\ \ \mathcal{M}'\left(0; 1, 1, n, \ell\right)$. Thus for a small monetary shock, $Std(\Delta p_i)$ has no effect on the cumulative output. The usefulness of the approach developed in this section is easily seen in the $n = 1$ case. Using the closed form solution for $g(p)$ in equation (3) and the expression for $m(p)$ in equation (14) we can analytically compute the cumulative effect of a small shock δ using equation (16) and the approximation $\mathcal{M}(\delta) \approx \delta\, \mathcal{M}'(0)$ which yields

$$\delta\, \mathcal{M}'(0) \quad = \quad \frac{\delta}{\epsilon\, N(\Delta p_i)}\ \frac{e^{\sqrt{2\phi}} + e^{-\sqrt{2\phi}}}{\left(e^{\sqrt{2\phi}} + e^{-\sqrt{2\phi}} - 2\right)^2} \left(e^{\sqrt{2\phi}} + e^{-\sqrt{2\phi}} - 2 - 2\phi\right)\ . \tag{18}$$

This formula can be used to compare a Calvo model where $\ell \to 1$ (or $\phi \to \infty$, see equation (6)) with a Golosov-Lucas model, where $\ell \to 0$ (or $\phi \to 0$). Simple analysis shows that in Calvo we have $\delta\,\mathcal{M}'(0) = \frac{\delta}{\epsilon\,N(\Delta p_i)}$, whereas in Golosov-Lucas we have $\delta\,\mathcal{M}'(0) = \frac{\delta}{6\epsilon\,N(\Delta p_i)}$, so that the cumulative output effect in these models differs by a factor of 6. Interestingly, the number 6 is exactly the ratio between the kurtosis of price changes in each of these models. The next proposition generalizes this result to any $n \geq 1$ and $\ell \in (0,1)$ (see Appendix B for the proof)

Proposition 6 *Consider an economy with $n > 1$, with steady-state statistics $N(\Delta p_i)$, $Std(\Delta p_i)$, ℓ and a steady-state kurtosis of price changes $Kur\,(\Delta p_i)$. For a small monetary shock δ we have the following first order Taylor expansion of $\mathcal{M}\,(\delta; N(\Delta p_i), Std(\Delta p_i), n, \ell)$:*

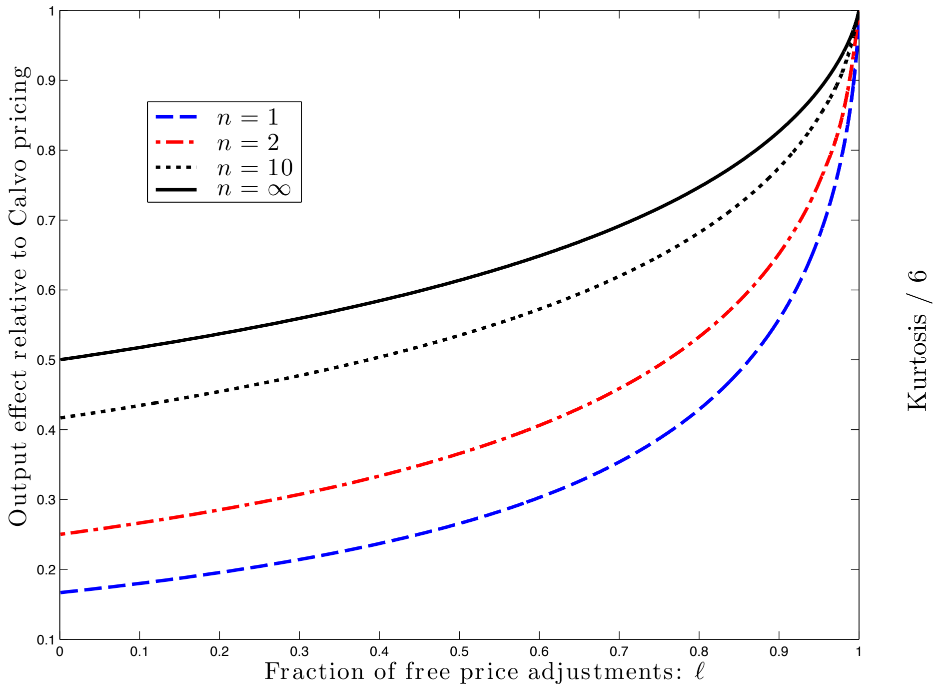
$$\delta\;\mathcal{M}'\,(0; N(\Delta p_i), Std(\Delta p_i), n, \ell) = \frac{\delta}{\epsilon}\,\frac{Kur\,(\Delta p_i)}{6\,N(\Delta p_i)} = \frac{\delta}{\epsilon}\,\frac{\sum_{i=2}^{\infty} \frac{1}{i!\,\Gamma\left(\frac{n}{2}+i\right)} \left(\frac{\lambda \bar{y}}{2\sigma^2}\right)^i}{\lambda \sum_{i=1}^{\infty} \frac{1}{i!\,\Gamma\left(\frac{n}{2}+i\right)} \left(\frac{\lambda \bar{y}}{2\sigma^2}\right)^i}\quad. \tag{19}$$

The proposition illustrates how it is possible for two models sharing similar features, e.g. calibrated to the same observables $N(\Delta p_i), Std(\Delta p_i)$, to have different output effects: what is needed is that the models feature a different kurtosis of price changes.

Recall from Proposition 4 that the shape of the size distribution of price changes, and hence kurtosis, depends only on n and ℓ. For a fixed n, kurtosis is increasing in ℓ. Indeed, as ℓ goes to 1 then kurtosis goes to 6, and hence we obtain $\mathcal{M}(\delta) \cong \delta/(\epsilon N(\Delta p_i))$, which is the result produced by the Calvo pricing model. On the other extreme, as ℓ goes to 0 kurtosis equals $3n/(n+2)$. This implies that, for instance, in the Golosov and Lucas case of $n = 1$, the impact of monetary policy is $1/6$ of Calvo. Also, keeping $\ell = 0$ and varying n from 1 to ∞, the effect goes from $1/6$ to $1/2$ of Calvo. Note that in the case of $\ell = 0$ and $n = \infty$ the model becomes Taylor's staggered price model or, equivalently, the Reis (2006) model. Thus the purely multi-product Taylor-Reis case ($\ell = 0$, $n = \infty$) delivers only half of the real effects compared to a purely Calvo model ($\ell = 1$), as further discussed below.

Figure 4 offers a richer systematic comparison of the real effects of monetary shocks as n and ℓ vary: the vertical axis plots the real output effect produced by a small monetary shock relative to the effect produced by a Calvo model where $\ell = 1$. Four curves are plotted in the figure, corresponding to $n = 1, 2, 10, \infty$. It appears that the model behavior for $n = 2$ remains quite close to the case where $n = 1$, as was also seen from the analysis of the distribution of price changes. Instead, the model behavior for $n = 10$ is quite close to that of a model where $n = \infty$. This is useful because the latter is quite tractable analytically, as discussed below. Figure 4 shows that at any level of ℓ the real output effects are smallest for $n = 1$. As explained in Alvarez and Lippi (2014) a larger number of goods dampens the selection effect, increasing the real output consequences of a monetary shock. Indeed at any level of ℓ

Figure 4: Cumulative output effect relative to Calvo pricing: $Kur(\Delta p_i)/6$



the effect is increasing in n. The figure shows that fixing n the output effect is increasing in ℓ. In the limit, as $\ell \to 1$ the economy converges to a Calvo model where the real effects are largest and independent of n.

The curves plotted in the figure are convex. In particular, some analysis reveals that the slope of the curve as $\ell \to 1$ diverges to $+\infty$ for any level of n.[17] The economic implication of this property is that a small deviation from Calvo pricing, i.e. a fraction of adjustment ℓ that is slightly below 1 is going to give rise to a large deviation from the real effects predicted by the Calvo pricing. That the relatively large real effects in Calvo are very sensitive to the introduction of a small amount of selection by firms regarding the timing of price changes is also apparent in the CalvoPlus model of Nakamura and Steinsson (2010) (see their figure VII). Hence the finding seems robust as these models, and their measures of real effects, are similar but not identical. Additionally, equation (19), together with equation (4) and equation (6), completes the closed form solutions for $N(\Delta p_i), Std(\Delta p_i), Kur(\Delta p_i)$ and $\mathcal{M}'(0)$ given parameters n, λ, σ^2 and threshold $\bar{y}$, which instead can be solved explicitly using Lemma 1 and Proposition 2.

We conclude by discussing the economics of why there is a systematic relationship between kurtosis and the cumulative output effect of a monetary shock. As noticed by Golosov and

[17]For a formal proof see the closed form solution for $\mathcal{M}$ in the $n = \infty$ case, in technical Appendix Gsec-2-closedform-cases.

Lucas (2007) the small output effect of a monetary shock in the menu cost model is related to the degree to which the firms that change prices right after the shock are of a "selected" type. Our result establishes that the steady state kurtosis of the price changes encodes the strength of such selection effects across several models. For instance in Golosov-Lucas the price changes in the steady state are concentrated at the value of the adjustment barrier (e.g. $\pm\bar{p}$), which yields the smallest possible value of kurtosis (equal to 1). This matters for the transmission of monetary shocks since the price changes immediately after a shock are going to be large, mostly of size $\Delta p_i = \bar{p}$, so that the CPI response is fast. By contrast, in Calvo the size of the average price change across adjusting firms (after a shock) equals the size of the shock, typically a small value δ (much smaller than the adjustment threshold $\bar{p}$). This happens from averaging across all small (a large mass) as well as large price changes, which occur due to the random choice of the adjusting firms, and are reflected in the high kurtosis (a peaked shape) of the size distribution of price changes in Calvo (equal to 6). This causes the CPI to rise more slowly and the real effects to be more persistent than in the Golosov-Lucas model.

Yet selection in the *size* of price adjustment is only one piece of the mechanism: the size of the price changes across adjusting firms in Calvo ($\ell = 1$) and in Taylor ($n = \infty, \ell = 0$) is exactly δ, and yet the cumulative output in Taylor is half of the cumulative output in Calvo. The difference across these models is in the *timing* of adjustments: in Taylor the time elapsed between adjustments is constant, while in Calvo it has an exponential distribution, with a thick right tail of firms that adjust very late (notice though that the expected time between adjustments coincides in the two models). This selection effect (in the times of price changes) is also captured by the steady state kurtosis, since a more dispersed distribution of times between adjustments produces a distribution of price changes that is a mixture of normals with different variances (recall the variance of the distribution is proportional to the time elapsed since the adjustment), whose kurtosis will be above the one of the normal distribution (see Lemma 2 below for a formalization). Indeed in Alvarez, Lippi, and Paciello (2016) we show that the result in Proposition 6 extends to a more general class of models where selection concerns exclusively the times of adjustment (not the size), as in the model of Carvalho and Schwartzman (2015).

Lack of sensitivity to inflation. While the model we have written is based on an economy which has zero steady-state inflation, we argue that the characterization of kurtosis as well as the key result for the effect of monetary policy in Proposition 6 apply also for low rates of inflation. Indeed steady-state inflation has only second-order effect on both the kurtosis of the price changes as well as the area under the IRF of output for a small monetary shock.

In particular, we show that both the left and right hand side of equation (19) have a zero derivative with respect to steady-state inflation, evaluated at zero inflation. For this we consider an economy where the money supply grows at the rate μ and the steady state inflation rate and the growth rate of nominal wages both equal μ. In this case the price gaps will evolve as $\mathrm{d}p_i(t) = -\mu \mathrm{d}t + \sigma \,\mathrm{d}W_i(t)$ where the negative sign of the drift reflects the fact that wages grow at a constant rate. For this model we let $Kur(\Delta p_i;\mu), N(\Delta p_i;\mu)$ be the kurtosis and frequency of price changes at steady state as a function of inflation μ. Likewise, we consider a once and for all (unanticipated) increase in the level of money supply of size δ, so that the path continues to grow at rate μ immediately afterwards. We let $\mathcal{M}(\delta;\mu)$ be the area-under the IRF of output after such monetary shock δ, also indexed by the steady state inflation. We have the following result (see Appendix B for the proof)

Proposition 7 *Let μ be the steady state inflation rate. Then: $0 = \frac{\partial\, Kur(\Delta p_i;\mu)}{\partial\, \mu}\Big|_{\mu=0} = \frac{\partial\, N(\Delta p_i;\mu)}{\partial\, \mu}\Big|_{\mu=0} = \frac{\partial\, \mathcal{M}'(0;\mu)}{\partial\, \mu}\Big|_{\mu=0}$, hence the derivative of both sides of equation (19) with respect to μ, evaluated at $\mu = 0$, is equal to zero.*

Hence, even though we have developed the result for zero inflation, this proposition shows that the results provide a good benchmark for a low inflation economy. We think this is important since developed economies have low but positive inflation rates. The idea behind the proof is to exploit the symmetry of both kurtosis $Kur(\Delta p_i;\mu)$ and the cumulative IRF $\mathcal{M}'(0;\mu)\delta$ with respect to inflation μ around $\mu = 0$.

To summarize, we showed that adding a small inflation has a negligible effect on the findings of the paper, including the main result in equation (1). This theoretical prediction is consistent with evidence on the small elasticity of several price setting statistics (such as the frequency, variance and kurtosis) provided in Gagnon (2009) and Alvarez et al. (2015).

4.2 The case of a large monetary shock

The dependence of the output effect on the size of the monetary shock is a hallmark of menu cost models. The next proposition characterizes the smallest value of the monetary shock $\underline{\delta}$ for which all firms adjust prices immediately, in which case the impact response of the aggregate price level is equal to the monetary shock, $\mathcal{P}(\delta, 0) = \delta$ so that prices are flexible and there is no effect on output, $\mathcal{M}(\delta) = 0$. The value of this threshold for the monetary shock is proportional to the $Std(\Delta p_i)$ which properly scales the "size" of the monetary shocks, and to a constant of proportionality that is increasing in ℓ (see Appendix B for the proof).

Proposition 8 *Define $\underline{\delta}$ as the smallest once-and-for-all monetary shock for which there is full price flexibility, i.e. $\mathcal{M}(\delta) = 0$ for any $\delta \geq \underline{\delta}$. We have $\underline{\delta} = 2\sqrt{\frac{\bar{y}}{n}} = 2\, Std(\Delta p_i)\sqrt{\frac{\mathcal{L}^{-1}(\ell;n)}{\ell}}$,* *where $\mathcal{L}(\cdot; n)$ is given by equation (6). Fixing any $n \geq 1$, the ratio $\underline{\delta}/(2Std(\Delta p_i))$ is a strictly increasing function of ℓ, ranging from 1 to ∞ as ℓ varies from 0 to 1.*

Note that for $\ell = 0$, we have $\underline{\delta} = 2\, Std(\Delta p_i)$ for all $n \geq 1$, so the minimum shock is simply twice the standard deviation of prices. In general this threshold is increasing in ℓ, becoming unbounded as the model gets closer to Calvo, i.e. $\lim_{\ell \to 1} \underline{\delta} = +\infty$ for all $n \geq 1$. In other words, $\underline{\delta}$ is a steep convex function of ℓ, with an infinite positive slope as $\ell \to 1$.

5 Robustness and scope of results

The class of models for which Proposition 6 holds is one where the innovations to the firm's costs are i.i.d. and normally distributed. In Section 5.1 we discuss the usefulness and realism of this assumption. In Section 5.2 we modify the model with normal innovations to the costs to one where cost innovations are drawn from a mixture of normals, and thus can display an arbitrarily large kurtosis. We start by establishing some time-aggregation properties of leptokurtic cost changes as a function of the time elapsed between observations $T > 0$. This property is important because as T grows large the distribution of the cost changes converges to normal, even though its innovations are leptokurtic, as an immediate consequence of the central limit theorem. Next, we parameterize the model using observations on the kurtosis of cost changes (i.e. wholesale prices) over time periods of different lengths. We then rederive a measure the real output effects of monetary shocks in this economy using the method developed in the previous section. The main result, summarized in Proposition 10 and illustrated in Figure 5, indicates that, even in the presence of cost shocks with large kurtosis (such as the estimated from wholesale price data), equation (1) remains informative to gauge the size of the real effects for actual economies.

5.1 Usefulness and realism of normally distributed cost shocks

Our analysis was built on the assumption that the (log) changes in costs are normally distributed. For instance in the firm level data on prices and costs reported in Figure 1 by Carlsson and Skans (2012) the distribution of the annual cost changes is close to normal, with a kurtosis coefficient of 3.8.[18] While the assumption of normally distributed cost shocks is stylized, it is particularly useful to evaluate whether price changes reflect cumulative small changes in costs. Indeed, the normal i.i.d. innovations of the costs is the only assumption needed for the result in Proposition 1, which shows that for almost any model

[18]Notice that data from all the 702 firms behind this statistic are pooled, i.e. not standardized, so that it is likely that the kurtosis is smaller.

$Var(\Delta p_i) \cdot N(\Delta p_i) = \sigma^2$. Thus, the variance of price changes is inversely proportional to the frequency of prices changes, for a given level of the innovations for cost σ^2. There is some support for this hypothesis, for instance Eichenbaum, Jaimovich, and Rebelo (2011): "*In our dataset prices are more volatile than our measure of marginal cost [...] The basic intuition for this result is that prices often do not change much in response to small changes in cost. But small cost changes that cumulate and create large deviations from the average markup can trigger very large changes in price.*"[19]

Another key assumption in deciding whether normally distributed cost shocks are realistic concerns the level of the production chain at which costs should be measured, or more broadly, the level of the production chain at which price stickiness is created. Should one measure the costs faced by retailers, or measure the costs at an early stage in the production-distribution chain? For instance, taking the perspective that "the price-setting firm" stands for the integrated production-wholesale-retail sector, where a non-negligible fraction of the cost is the price of raw materials and/or imported inputs, the assumption that costs follow a random walk with normally distributed innovations is reasonable.[20] Some evidence consistent with this perspective is provided by Nakamura and Zerom (2010) and by Goldberg and Hellerstein (2013) who conduct two interesting case studies of coffee and imported beer, respectively. They find that the frequency of price changes at the retail level is essentially the same as the one for wholesale prices changes faced by the retailers, and that there is a very high pass-through between wholesale and retail prices–see Figures 1 and 2 in both of these papers. For these products, changes in the price of coffee beans and in the exchange rates –both realistically modeled as random walks with normal innovations– are important determinants of the cost at an earlier stage of production. Thus it appears that for these products the stickiness was generated before the last retail stage. To complement these detailed case studies we used a data set with prices charged by producers to wholesalers ("PromoData") with broad coverage both geographically and in terms of the number of products. We compare the frequencies of price changes from "PromoData" with the frequencies from a retailer (scanner) data of very broad coverage (the IRI Symphony). We find that prices charged by producers to wholesalers are at least as sticky as prices charged by retailers to households –see Appendix F for more details. Nevertheless, establishing which is the more reasonable level of aggregation and properly measuring them is a hard question which we leave for future research. Hence, to be conservative, in the rest of this section we take the narrow perspective

[19]For instance, the bottom panel of Table 2 in their paper estimates that the standard deviation of the Log changes in weekly prices, conditional on a price change, is 0.20. This is considerably higher than the standard deviation of Log changes in weekly cost, estimated to be 0.11 – or in the case of reference prices and reference cost of 0.14 and 0.06.

[20]Additionally, from the perspective of an integrated sector, one can obtain normality of the cost shocks by combining a large number of inputs with independent shocks, by virtue of the central limit theorem.

of the price setting firm as a retailer.

5.2 A menu-cost model with leptokurtic cost-shocks

Let $x(t)$ measure the change of the firm's marginal cost at time t, which obeys:

$$dx(t) = \sigma(t)\; dW(t) \;\; \text{with } x(0) = 0 \;\; \text{and } x(T) = \int_0^T \sigma(t)\, dW(t) \tag{20}$$

where $\{W(t)\}$ is a standard Brownian motion, and where $\{\sigma(t)\}$ is a process independent of $\{W(t)\}$. We are interested in the kurtosis of the cost changes over a period of length T, $K(T) = \frac{\mathbb{E}_0\left[x(T)^4\right]}{(\mathbb{E}_0[x(T)^2])^2}$. We obtain the following lemma (see Appendix B for the proof)

Lemma 2 *Assume that $\sigma(t)$ is uniformly bounded from above we have:*

$$K(T) = 6\;\frac{\int_0^T \left[\int_0^t \mathbb{E}_0\left[\,\sigma(t)^2\,\sigma(s)^2\,\right]\, ds\right]\, dt}{\left(\int_0^T \mathbb{E}_0\left[\sigma(t)^2\right]\, dt\right)^2} \tag{21}$$

Inspection of equation (21) shows that when x is a standard Brownian motion, so that $\sigma(t) = \bar{\sigma}$, we have that $K(T) = 6\;\frac{\bar{\sigma}^4\, T^2/2}{\bar{\sigma}^4\, T^2} = 3$, which is obvious since $x(T)$ is normally distributed at all horizons T. Instead, if the changes in cost shocks are caused by innovations with different volatility then the mixture of normals generated by equation (20) may give rise to arbitrarily large values of excess Kurtosis, depending on the behavior of $\{\sigma(t)\}$ and the time interval T over which the change in the cost is measured.

In what follows, we specialize $\{\sigma(t)\}$ to be a continuous-time two-state Markov chain. We assume that in one state the cost innovations have a low variance $\sigma_0 \geq 0$, and that in the other state they have a larger variance $\sigma_1 > \sigma_0$. In particular we assume that there is a state $u(t) \in \{0, 1\}$, and that if $u(t) = i$ then the state changes to $u(t + dt) = j \neq i$ with a probability $\theta_i\; dt$ where $i, j \in \{0, 1\}$. Then for all t we have: $\sigma(t) = \sigma_{u(t)}$, so there are four non-negative parameters describing the problem, the two exit rates θ_0, θ_1 and the two values for the instantaneous volatility: σ_0, σ_1. This simple specification of the process for cost shocks allows us to approximate several menu cost models, considered in the literature, which deviate from the assumption of normal shocks. For instance Gertler and Leahy (2008) or Midrigan (2011) can be modeled as $\sigma_0 \to 0$ and a fixed rate of arrival (θ_0) of the large shocks (parametrized by σ_1, whose persistence is given by θ_1). Likewise the setup can be used to study the discrete-time model by Karadi and Reiff (2014) in which the weekly innovations to cost shocks are a mixture of two normal distributions (characterized by σ_0 and σ_1) and the relative frequency and persistence of each of these two states (characterized by θ_0 and θ_1).

The next proposition characterizes the kurtosis $K(T)$ given these 4 fundamental parameters (see Appendix B for the proof). We have:

Proposition 9 *Assume that the volatility follows a two-state continuous time Markov process as described above. Without loss of generality parameterize the process for x in terms of the average rate of change of the state, θ, the fraction of time spent in state 1, s, and the ratio of the standard deviations of the two states, ξ:*

$$\theta \equiv \frac{1}{2}(\theta_1 + \theta_0) \ , \ s \equiv \frac{\theta_0}{\theta_0 + \theta_1} \ , \quad \text{and } \xi \equiv \frac{\sigma_0}{\sigma_1}$$

Then the kurtosis of the cumulative changes up to horizon $T \geq 0$ is given by

$$K(T; \theta, \xi, s) = 3 + 6 \, \frac{(1-\xi^2)^2 \, s(1-s)}{[\xi^2(1-s) \, + \, s]^2} \, \frac{\left[2\,T\,\theta - 1 + e^{-2\,T\,\theta}\right]}{(2\,T\,\theta)^2} \tag{22}$$

This proposition shows that the kurtosis of the cost shocks is monotonically decreasing in $T\theta$, converging towards the value of 3 as $T\theta$ becomes large. The latter feature is easily understood: as the time period become sufficiently long (large T) or as the switching rate between states with high and low variance increases (high θ) then the mixture of normals converges towards the normal distribution (a weighted sum of normals).

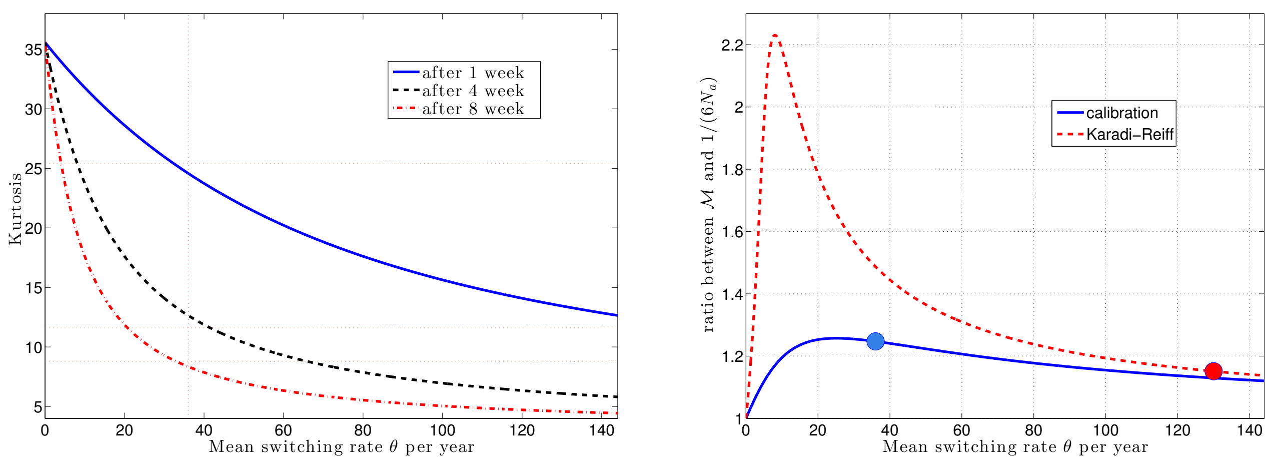
We consider three values of T, corresponding to empirical observations of the cost changes after 1-week ($T = 1/52$), 4-weeks ($T = 4/52$) and 8-weeks ($T = 8/52$). Our interest in these 3 moments is motivated by the corresponding moments of cost changes observed in the Dominick's dataset, previously used in the sticky price literature.[21] In what follows we present results obtained with a minimal filtering of the cost data, which yields much higher values for kurtosis (sensitive to extreme observations). We see this case as useful for the robustness analysis, even though a more reasonable analysis might involve more trimming of the tails of the cost distributions and some weighting of the different UPC, producing results that are even closer to the assumption of Normally distributed shocks. The data show that the kurtosis of cost changes measured over 1-week is 25.4, over 4-weeks is 11.6 and over 8-weeks is 8.8. Note that these estimates include the cost changes equal to zero (those obtained when the level

[21] Following Midrigan (2011) we focus on store #122 (see his Table 8), the one with most observations. For each observation, the cost is measured by: cost=price*(100-profit margin)/100. The change in cost is 100 times the log-change in cost over 1, 4 and 8 weeks. We apply the following trimming to the data: we drop observations whose cost is smaller than 20 cents or larger than 25 dollars (deemed implausible in view of the type of items sold and the distribution of price levels); we drop non-zero cost changes that are smaller than 1 cent (in absolute value); we drop as outliers observations with log-cost changes larger in absolute value than the 99^{th} percentile of absolute log cost changes. The resulting number of observations of cost changes is 946,315 for 1 week changes, 924,363 for 4-week changes, and 900,783 for 8-week changes. As for price changes, we standardize cost changes at the UPC level (around 15,000 UPCs are available). The data is publicly available at http://research.chicagobooth.edu/kilts/marketing-databases/dominicks.

of the cost remains constant), which might contribute to a high level of kurtosis. It appears that kurtosis is decreasing in T, consistently with the central prediction of Proposition 9: as more time elapses more averaging occurs (between low and high variance states), and thus kurtosis decreases. Given the 3 values of T corresponding to these 3 moments, we pin down the remaining 3 parameters, $\{\xi, s, \theta\}$ by matching the data. The parameters underlying this calibration imply a ratio between the standard deviation $\xi = \sigma_0/\sigma_1 = 0.15$, the mass of firms with high variance is $s = 0.022$ and that $\theta = (\theta_0+\theta_1)/2 = 36$. The left panel of Figure 5 plots the theoretical lines associated to each T at the calibrated values of ξ and s as a function of θ. The three observed moments are marked by thin horizontal lines. It appears that at $\theta = 30$ the theory is essentially on top of the data.

Figure 5: Model with leptokurtic cost shocks ($n = 1, \ell = 0$)

Kurtosis of cost changes after T-weeks ($T = 1, 4, 8$) Cumulative output effect relative to $\frac{1}{6N_a}$



Three dashed horizontal lines (left panel) indicate the kurtosis values observed in the data over the 1-, 4-, and 8-week durations (Dominick's data). The calibration of the Kurtosis of cost shocks uses $\xi = 0.14$, the mass of firms with high variance is $s = 0.022$ and that $\theta = (\theta_0 + \theta_1)/2 = 36$. The right panel uses these calibrated parameters and $\bar{p}/\sigma_1$ such that there is one price adjustment per year, $N_a = 1$.

Real output effects. Next, we apply the analytical methods developed in Section 4 to compute $\mathcal{M}(\delta)$, the expected cumulated output effect for a small monetary shock δ to this environment with leptokurtic cost shocks. For simplicity, we consider a simple price setting problem where the firm sells a single good, $n = 1$, and there are no free-adjustment opportunities $\ell = \lambda = 0$. We assume that firms follow a threshold rule such that the price is adjusted if the price gap falls outside the inaction interval $(-\bar{p}, \bar{p})$.[22] Under normally distributed cost shocks these assumption correspond to the Golosov and Lucas model where the kurtosis of price changes equals 1 and the cumulated output effect is equal to $\mathcal{M}(\delta) = 1/(6N_a)$, as from Proposition 6. Our objective is to analyze what is the cumulated output effects in this model once the shocks are leptokurtic.

We use the definition of cumulative output gap after a shock of size δ to define the cumulated real output effect as:

$$\mathcal{M}(\delta) = \int_{-\bar{p}}^{\bar{p}-\delta} \Big(m_0(p)\, g_0(p+\delta) + m_1(p)\, g_1(p+\delta) \Big)\, dp$$

which is the analogue of equation (16) in this model with a single good ($n = 1$), no free-adjustment opportunities $\ell = 0$ and a two-state Markov process for the variance of the cost shocks. The key novelty compared to Section 4 is that there are now two sets of firms (one for each high/low variance state), as indicated by the densities $g_i(p)$, and two measures of expected cumulated output associated to each firm in each state, $m_i(p)$. The next proposition characterizes the frequency of price adjustment N_a and the real output effect $\mathcal{M}(\delta)$ in terms of few fundamental parameters (see Appendix B for the proof and closed form expressions for $g_i(p)$ and $m_i(p)$):

Proposition 10 *Fix the 3 parameters $\{\theta, s, \xi\}$ that pin down $K(T)$, and fix the ratio $\bar{p}/\sigma_1$. The mean time between price adjustments is:*

$$\frac{1}{N_a} = \frac{\bar{p}^2/\sigma_1^2}{s\,(\xi^2\hat{\rho}+1)} + \frac{\hat{\rho}(1-\xi^2)^2}{2\theta(\xi^2\hat{\rho}+1)^2}\left(1 - \frac{2}{e^{\chi\bar{p}} + e^{-\chi\bar{p}}}\right) \tag{23}$$

$$\textit{where} \quad \hat{\rho} \equiv \frac{\theta_1}{\theta_0} = \frac{1-s}{s} \quad \textit{and} \quad \chi \equiv \sqrt{\frac{4\theta s(\xi^2\hat{\rho}+1)}{\sigma_1^2\,\xi^2}}$$

The cumulated output following a small monetary shock of size δ is

$$\mathcal{M}(\delta) \approx \delta\ \mathcal{M}'(0) = \frac{\delta}{\epsilon}\left[\frac{\bar{p}^2/\sigma_1^2}{6(1+\xi^2\hat{\rho})s} - \frac{\hat{\rho}(1-\xi^2)^2}{\theta(1+\xi^2\hat{\rho})^2}\left(\frac{e^{\chi\bar{p}} + e^{-\chi\bar{p}} - 2}{\chi\bar{p}(e^{\chi\bar{p}} - e^{-\chi\bar{p}})} - \frac{1}{2}\right)\right] \quad . \tag{24}$$

The proposition states that the frequency of adjustment N_a is a function of 4 fundamental parameters. Three parameters are pinned down by the data on kurtosis, $\{\theta, \xi, s\}$, a new parameter $\bar{p}/\sigma_1$ is pinned down to match the data on N_a. Notice that given these 4 parameters the cumulated real effects of a small monetary shock $\mathcal{M}(\delta)$ in equation (24) is completely

[22] We follow this rule for comparability with the literature. The optimal threshold rule $\bar{p}_{u(t)}$ is in general indexed by the state. A full solution for the latter case is available from the authors.

determined. This result is important because it identifies the determinants of the cumulated real effects of a small monetary shock. The right panel in Figure 5 plots the cumulated output effects $\mathcal{M}(\delta)$ relative to the same effect produced by the model with Normal cost shocks (equal to $1/(6N_a)$) as a function of θ, the arithmetic average of the number of state switches per year. Some analysis shows that equation (24) equals 1 as $\theta \to 0$ or as $\theta \to \infty$, independent of all other parameters. The reason is that under these extreme cases the distribution of cost shocks is again Normal, and the real effects coincide with the ones in Golosov and Lucas (so the ratio of the effects is 1). At intermediate values of θ the cumulated output effect is bigger in the model with leptokurtic shocks, and the exact amount depends on the parameterization. The right panel of the figure plots two curves, both consistent with a frequency of 1 price adjustment per year, $N_a = 1$. The lower thick curve uses the parameters $\{\xi, s\}$ that were chosen to calibrate the data on kurtosis. This corresponds to our "preferred" specification since it is the one that matches the cost data more closely. At these values the figure shows that the cumulated real effect is about 25% larger than the effect that is obtained in the model with normally distributed shocks (marked by a thick dot in the figure). The higher dashed curve uses values of ξ and s that are implied by the parametrization of Karadi and Reiff (2014).[23] These values imply a kurtosis of cost changes that is smaller than the one in the data (namely of 12, 6 and 5, over the 1-, 4- and 8-week duration).[24] Nevertheless we consider this case as a robustness check. The figure shows that, using the parametrization of Karadi and Reiff, the cumulated real effect is close to what these authors report in their paper (see the bold bullet on the dashed curve), and aligned to the values predicted by our calibration, only 25% above the value predicted by Proposition 6. To put this approximation error into perspective, consider that the order of magnitudes involved in the discussion on the size of the real effects of monetary shocks in sticky price models are in the order of 600%, as given by the ratio between the real cumulative effect in a Calvo model ($\ell = 1$) versus the real effect in a Golosov-Lucas menu cost model ($\ell = 0, n = 1$).

Finally, we notice that we developed our robustness analysis using a model with a single good, $n = 1$. Considering multi-product firms would further strengthen the robustness of our findings. The reason is easily seen: with 1 good the firm is "averaging" high and low variance shocks across time, so that the cost shock innovations become normally distributed as $T\theta$ becomes large (see equation (22)). As n increases the approximation error shrinks further

[23] See their Table 2. Translating their weekly parameters into ours gives $\xi = 0.088$ for the ratio of standard deviations, $1 - s = 0.912$ or that 91.2% of firms are in state 0. They assume that the probability per week of drawing a high-variance state is $p = 0.912$, iid distributed. This implies an expected duration in weeks of $p/(1-p) + p/(1-p)^2$. Once this is translated into duration in years it is immediate to map it into our θ_i parameter. Using their value for p we get $\theta_0 = 0.40$ and $\theta_1 = 257$, or $\theta \cong 129$.

[24] Instead, a parametrization as in Midrigan or Gertler and Leahy, in which $\xi = 0$, predicts a value of kurtosis at the 1 week frequency of around 80, much higher than in the data.

since the firm's cost shocks are now averaged both across time as well as across goods (at a given point in time).[25]

6 Concluding remarks

This paper develops an analytical model that is able to match the cross sectional patterns on the frequency, variance and kurtosis of price changes, for small values of the (menu) cost of price adjustment. Our model nests several classic models of price setting, such as Taylor (1980); Calvo (1983); Reis (2006); Golosov and Lucas (2007); Midrigan (2011); Nakamura and Steinsson (2010); Alvarez and Lippi (2014), and sheds light on the propagation of monetary shocks. The main finding is that the real cumulative output effect of a monetary shock is proportional to the kurtosis of the size-distribution of price changes. The sizable differences produced by previous models can be largely explained in terms of their different predictions for the kurtosis of price changes.

Our main result, namely equation (1), is robust to realistic changes in the setup, such as a small inflation rate, and some (data consistent) degree of non-normality of the cost shocks. The result also holds in a version of our model, closer to the original CalvoPlus model by Nakamura and Steinsson (2010), which assumes that the price adjustment upon a random opportunity is cheaper than the regular menu cost but not free (see the technical Appendix U). This result also emerges in different model setups: in Alvarez, Lippi, and Paciello (2016) we consider the case where firms face a random observation cost and no menu costs. This setup produces random adjustment times at the firm level, whose consequences for the propagation of monetary shocks have been studied by Mankiw and Reis (2002) and Carvalho and Schwartzman (2015). The last class of models does not feature any selection in the size of price changes but only in the times of adjustment.[26] Interestingly, one can show analytically that equation (1) holds in this class of models, too. More broadly, in Alvarez, Lippi, and Paciello (2011) we analyzed the optimal decision problem of a firm that faces both an observation as well as a menu cost – so that the decision rules of both Reis (2006) and Golosov and Lucas (2007) are obtained as special cases. Numerical simulations for the aggregate economy implied by this setup suggest that equation (1) continues to hold. In Alvarez, Lippi, and Passadore (2016) we discuss differences and similarities of models featuring time dependent versus state dependent rules. We conclude that for small aggregate shocks equation (1) continues to hold irrespective of the fundamental friction that causes

[25] Technically the state variable for a multi product firm is the sum of the (squared) price gaps, whose law of motion follows the Bessel process (the sum of n squared Brownian motions). In the presence of leptokurtic shocks convergence to this law of motion is faster as n increases.

[26] Moreover, while kurtosis is between 1 and 6 in our menu cost model, it is unbounded in these models.

price stickiness.

References

Alvarez, Fernando, Martin Gonzalez-Rozada, Andres Neumeyer, and Martin Beraja. 2015. "From Hyperinflation to Stable Prices: Argentina's evidence on menu cost models." manuscript, University of Chicago.

Alvarez, Fernando, Francesco Lippi, and Juan Passadore. 2016. "Are State and Time dependent models really different?" *NBER Macroeconomics Annual 2016, Volume 31* .

Alvarez, Fernando E. and Francesco Lippi. 2014. "Price setting with menu costs for multi product firms." *Econometrica* 82 (1):89–135.

Alvarez, Fernando E., Francesco Lippi, and Luigi Paciello. 2011. "Optimal price setting with observation and menu costs." *The Quarterly Journal of Economics* 126 (4):1909–1960.

———. 2016. "Monetary Shocks in Models with Inattentive Producers." *Review of Economic Studies* 83:421–459.

Alvarez, Luis J., Emmanuel Dhyne, Marco Hoeberichts, Claudia Kwapil, Herve Le Bihan, Patrick Lunnemann, Fernando Martins, Roberto Sabbatini, Harald Stahl, Philip Vermeulen, and Jouko Vilmunen. 2006. "Sticky Prices in the Euro Area: A Summary of New Micro-Evidence." *Journal of the European Economic Association* 4 (2-3):575–584.

Arkolakis, Costas, Arnaud Costinot, and Andres Rodriguez-Clare. 2012. "New Trade Models, Same Old Gains?" *American Economic Review* 102 (1):94–130.

Barro, Robert J. 1972. "A Theory of Monopolistic Price Adjustment." *Review of Economic Studies* 39 (1):17–26.

Berardi, Nicoletta, Erwan Gautier, and Hervé Le Bihan. 2015. "More Facts about Prices: France Before and During the Great Recession." *Journal of Money, Credit and Banking* 47 (8):1465–1502.

Bhattarai, Saroj and Raphael Schoenle. 2014. "Multiproduct Firms and Price-Setting: Theory and Evidence from U.S. Producer Prices." *Journal of Monetary Economics* 66:178–192.

Bils, Mark and Peter J. Klenow. 2004. "Some Evidence on the Importance of Sticky Prices." *Journal of Political Economy* 112 (5):947–985.

Bonomo, Marco, Carlos Carvalho, and Ren Garcia. 2010. "State-Dependent Pricing under Infrequent Information: A Unified Framework." Tech. rep., SSRN.

Calvo, Guillermo A. 1983. "Staggered prices in a utility-maximizing framework." *Journal of Monetary Economics* 12 (3):383–398.

Carlsson, Mikael and Oskar Nordstrom Skans. 2012. "Evaluating Microfoundations for Aggregate Price Rigidities: Evidence from Matched Firm-Level Data on Product Prices and Unit Labor Cost." *American Economic Review* 102 (4):1571–95.

Carvalho, Carlos. 2006. "Heterogeneity in Price Stickiness and the Real Effects of Monetary Shocks." *The B.E. Journal of Macroeconomics* 6 (3):1–58.

Carvalho, Carlos and Felipe Schwartzman. 2015. "Selection and Monetary Non-Neutrality in Time-Dependent Pricing Models." *The Journal of Monetary Economics* (forthcoming).

Cavallo, Alberto. 2015. "Scraped Data and Sticky Prices." NBER Working Papers 21490, National Bureau of Economic Research, Inc.

Cavallo, Alberto, Brent Neiman, and Roberto Rigobon. 2014. "Currency Unions, Product Introductions, and the Real Exchange Rate." *The Quarterly Journal of Economics* 129 (2):529–595.

Cavallo, Alberto and Roberto Rigobon. 2016. "The Billion Prices Project: Using online prices for measurement and research ." Working paper 22111, NBER.

Chakrabarti, Rajesh and Barry Scholnick. 2007. "The mechanics of price adjustment: new evidence on the (un) importance of menu costs." *Managerial and Decision Economics* 28 (7):657–668.

Chetty, Raj. 2009. "Sufficient Statistics for Welfare Analysis: A Bridge Between Structural and Reduced-Form Methods." *Annual Review of Economics* 1 (1):451–488.

Christopoulou, Rebekka and Philip Vermeulen. 2012. "Markups in the Euro area and the US over the period 19812004: a comparison of 50 sectors." *Empirical Economics* 42 (1):53–77.

Dixit, Avinash. 1991. "Analytical Approximations in Models of Hysteresis." *Review of Economic Studies* 58 (1):141–51.

Dotsey, Michael, Robert G. King, and Alexander L. Wolman. 2009. "Inflation and Real Activity with firm-level productivity shocks." Tech. rep., Federal Reserve Bank of Richmond.

Dutta, Shantanu, Mark Bergen, Daniel Levy, and Robert Venable. 1999. "Menu Costs, posted prices and multiproduct retailers." *The Journal of Money, Credit and Banking* 31 (4):683–703.

Eichenbaum, Martin, Nir Jaimovich, and Sergio Rebelo. 2011. "Reference Prices, Costs, and Nominal Rigidities." *American Economic Review* 101 (1):234–62.

Eichenbaum, Martin S., Nir Jaimovich, Sergio Rebelo, and Josephine Smith. 2014. "How Frequent Are Small Price Changes?" *American Economic Journal: Macroeconomics* 6 (2):137–155.

Gagnon, Etienne. 2009. "Price Setting During Low and High Inflation: Evidence from Mexico." *Quarterly Journal of Economics* 124 (3):1221–1263.

Gertler, Mark and John Leahy. 2008. "A Phillips Curve with an Ss Foundation." *Journal of Political Economy* 116 (3):533–572.

Goldberg, Pinelopi Koujianou and Rebecca Hellerstein. 2013. "A Structural Approach to Identifying the Sources of Local Currency Price Stability." *Review of Economic Studies* 80 (1):175–210.

Golosov, Mikhail and Robert E. Jr. Lucas. 2007. "Menu Costs and Phillips Curves." *Journal of Political Economy* 115:171–199.

Karadi, Peter and Adam Reiff. 2014. "Menu Costs, Aggregate Fluctuations, and Large Shocks." Working Paper Series 10138, CEPR.

Kashyap, Anil K. 1995. "Sticky Prices: New Evidence from Retail Catalogs." *The Quarterly Journal of Economics* 110 (1):245–74.

Kiley, Michael. 2002. "Partial Adjustment and Staggered Price Setting." *Journal of Money, Credit and Banking* 34 (2):283–98.

Klenow, J. Peter and Benjamin Malin. 2010. "Microeconomic Evidence on Price-Setting." *Handbook of monetary economics* 3:231–284.

Klenow, Peter J. and Oleksiy Kryvtsov. 2008. "State-Dependent or Time-Dependent Pricing: Does It Matter for Recent U.S. Inflation?" *The Quarterly Journal of Economics* 123 (3):863–904.

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

Mankiw, N. Gregory and Ricardo 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.

Midrigan, Virgiliu. 2011. "Menu Costs, Multi-Product Firms, and Aggregate Fluctuations." *Econometrica,* 79 (4):1139–1180.

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

———. 2010. "Monetary Non-Neutrality in a Multisector Menu Cost Model." *The Quarterly Journal of Economics* 125 (3):961–1013.

Nakamura, Emi and Dawit Zerom. 2010. "Accounting for Incomplete Pass-Through." *Review of Economic Studies* 77 (3):1192–1230.

Reis, Ricardo. 2006. "Inattentive producers." *Review of Economic Studies* 73 (3):793–821.

Sheedy, Kevin D. 2010. "Intrinsic inflation persistence." *Journal of Monetary Economics* 57 (8):1049–1061.

Taylor, John B. 1980. "Aggregate Dynamics and Staggered Contracts." *Journal of Political Economy* 88 (1):1–23.

Vavra, Joseph. 2014. "Inflation Dynamics and Time-Varying Volatility: New Evidence and an Ss Interpretation." *The Quarterly Journal of Economics* 129 (1):215–58.

Woodford, Michael. 2009. "Information-Constrained State-Dependent Pricing." *Journal of Monetary Economics* 56:s100–s124.

Wulfsberg, F. 2010. "Price adjustments and inflation-evidence from Norwegian consumer price data 1975-2004." Working paper, Bank of Norway.

Zaitsev, V.F. and A.D. Polyanin. 2003. *Handbook of Exact Solutions for Ordinary Differential Equations.* Chapman and Hall, CRC Press.

Zbaracki, Mark J., Mark Ritson, Daniel Levy, Shantanu Dutta, and Mark Bergen. 2004. "Managerial and Customer Costs of Price Adjustment: Direct Evidence from Industrial Markets." *The Review of Economics and Statistics* 86 (2):514–533.

APPENDICES – FOR ONLINE PUBLICATION ONLY

A Comparing BBP to CPI data to estimate kurtosis

We match a subset of our French CPI data with the prices from 3 French retailers taken from the Billion Price Project (BPP) dataset, see Cavallo (2015).[27] Table 2 offers two comparisons. The first three columns compare the BPP data from 2 large retailers with our CPI data for a similar type of outlet: to this end we restrict our dataset to CPI price records in "hypermarkets", excluding gasoline. The last two columns compare the BPP data from a large retailer specialized in electronics and appliances with the CPI data for goods in the category of appliances and electronic (we use the Coicop nomenclature, collected in outlets type "hypermarkets","supermarkets", and "large area specialists"). Comparing the values of kurtosis from both data sets suggests that $\Omega/\zeta \cong 2$, see equation (2). We can apply this magnitude to the full sample of CPI data, for which no "measurement error-free" counterpart like the BPP exists (and the feasible correction for heterogeneity is only partial), to obtain a corrected kurtosis. The number thus obtained for the kurtosis is near 4, so it lays in between the kurtosis of the Normal and the Laplace distribution.

Table 2: Comparison of the CPI vs. the BPP data in France

CPI category:	Hypermarkets			Appliances and electronic	
Data source:	BPP retailer 1	BPP retailer 5	CPI Hypermarkets	BPP retailer 4	CPI Large ret. electr.
duration (months)	8.6	8.1	4.8	6.4	7.2
kurtosis	5.5	4.3	10.1	2.8	6.3

Note: The BPP data are documented in Cavallo (2015). Results were communicated by the author. For CPI data source is INSEE, monthly price records from French CPI, data from 2003:4 to 2011:4. The sub-sample in the third column features the CPI records for the outlet type "hypermarkets". The sub-sample in the 5th column features the CPI records in the category of "appliances and electronic", as identified using the Coicop nomenclature, collected in the following outlets type: "hypermarkets","supermarkets", and "large area specialists". Data are standardized within each subsample using Coicop categories.

B Proofs

Proof. (of Proposition 1). Let $p(0) = 0$. Define $x(t) \equiv ||p(t)||^2 - n\,\sigma^2\,t$ for $t \geq 0$. Using Ito's lemma we can verify that the drift of $||p||^2$ is $n\sigma^2$, and hence $x(t)$ is a Martingale. By

[27] We are extremely grateful to Alberto Cavallo for sharing part of his data with us.

the optional sampling theorem $x(\tau)$, the process stopped at τ, is also a martingale. Then $\mathbb{E}\left[x(\tau) \;\middle|\; p(0)\right] = \mathbb{E}\left[||p(\tau)||^2 \;\middle|\; p(0)\right] - n\sigma^2 \mathbb{E}\left[\tau \;\middle|\; p(0)\right] = x(0) = 0$ and since $N(\Delta p_i) = 1/\mathbb{E}\left[\tau \;\middle|\; p(0)\right]$ and $Var(\Delta p_i) = \mathbb{E}\left[||p(\tau)||^2 \;\middle|\; p(0)\right]/n$ we obtain the desired result. □

Proof. (of Lemma 1). First, note that since two value functions differ by a constant, then all their derivatives are identical. Hence, if the one for the discount rate and arrival rate of free adjustment $(r+\lambda, 0)$ satisfies value matching and smooth pasting, so does the one for discount rate and arrival rate of free adjustment $(r, \lambda, 0)$, for the same boundary. Second, consider the range of inaction, subtracting the value function for the problem with parameters $(r+\lambda, 0)$ from the one with parameters (r, λ), and using that all the derivatives are identical, one verifies that if the Bellman equation holds for the problem with $(r+\lambda, 0)$, so it does for the problem with (r, λ). □

Proof. (of Proposition 2). The first part is straightforward given Lemma 1 and Proposition 3 in Alvarez and Lippi (2014). The second part is derived from the following implicit expression determining $\bar{y}$ (see the proof of Proposition 3 in Alvarez Lippi for the derivation):

$$\psi \;=\; \frac{B}{r+\lambda}\bar{y}\left[1 - \frac{\frac{2\sigma^2(n+2)}{r+\lambda}\bar{y} + \bar{y}^2 + \bar{y}^2 \sum_{i=1}^{\infty} \kappa_i \; (r+\lambda)^i \; \bar{y}^i}{\frac{2\sigma^2(n+2)}{r+\lambda}\bar{y} + 2\bar{y}^2 + \bar{y}^2 \sum_{i=1}^{\infty} \kappa_i \; (i+2) \; (r+\lambda)^i \; \bar{y}^i}\right] \tag{25}$$

where $\kappa_i = (r+\lambda)^{-i} \prod_{s=1}^{i} \frac{1}{\sigma^2(s+2)(n+2s+2)}$. So we can rewrite equation (25) as: $\psi = \frac{B}{r+\lambda}\bar{y}$ $\left[1 - \xi(\sigma^2, r+\lambda, n, \bar{y})\right]$. Since $\bar{y} \to \infty$ as $\psi \to \infty$ then we can define the limit:

$$\lim_{\psi \to \infty} \frac{\psi}{\bar{y}} = \frac{B}{r+\lambda}\left[1 - \lim_{\bar{y} \to \infty} \xi(\sigma^2, r+\lambda, n, \bar{y})\right]$$

Simple analysis can be used to show that $\lim_{\bar{y} \to \infty} \xi(\sigma^2, r+\lambda, n, \bar{y}) = 0$ which gives the expression in the proposition (see the technical Appendix J for a detailed derivation). □

Proof. (of Proposition 3). To characterize $N(\Delta p_i)$ we write the Kolmogorov backward equation for the expected time between adjustments $\mathcal{T}(y)$ which solves: $\lambda \mathcal{T}(y) = 1 + \; n\,\sigma^2\, \mathcal{T}'(y) + \; 2\, y\, \sigma^2\, \mathcal{T}''(y)$ for $y \in (0, \bar{y})$ and $\mathcal{T}(\bar{y}) = 0$ (see the technical Appendix K for details on the solution to this equation). Then the expected number of adjustments is given by $N(\Delta p_i) = 1/\mathcal{T}(0)$, subject to $\mathcal{T}(0) < \infty$.

The solution of the second order ODE for $\mathcal{T}(y)$ has a power series representation: $\mathcal{T}(y) = \sum_{i=0}^{\infty} \; \alpha_i \; y^i$, for $y \in [0, \bar{y}]$, with the following conditions on its coefficients $\{\alpha_i\}$: $\alpha_1 = \frac{\lambda\alpha_0 - 1}{n\sigma^2}$, $\alpha_{i+1} = \frac{\lambda}{(i+1)\;\sigma^2\;(n+2i)}\;\alpha_i$, for $i \geq 1$ and where $0 < \alpha_0 < 1/\lambda$ is chosen so that $0 \geq \alpha_i$ for $i \geq 1$, $\lim_{i \to \infty} \frac{\alpha_{i+1}}{\alpha_i} = 0$ and $0 = \sum_{i=0}^{\infty} \; \alpha_i \; \bar{y}^i$. Moreover, $\mathcal{T}(0) = \alpha_0$ is an increasing function of $\bar{y}$ since α_0 solves:

$$0 = \alpha_0 + \frac{(\alpha_0 - 1/\lambda)}{n} \left(\frac{\bar{y}\lambda}{\sigma^2}\right) \left[1 + \sum_{i=1}^{\infty} \left(\prod_{k=1}^{i} \frac{1}{(k+1)\ (n+2k)}\right) \left(\frac{\bar{y}\lambda}{\sigma^2}\right)^i\right]$$

Note that for $i \geq 1$: $\alpha_i = \alpha_i / [i!\ (n/2 + i)] (\lambda/(2\,\sigma^2))$, and using the properties of the Γ function

$$\alpha_i = \Gamma(n/2)/ \left(\Gamma(n/2+i) \left(\lambda/(2\,\sigma^2)\right)^i (\alpha_0 - 1/\lambda)\right. .$$

Solving for α_0 and using $\mathcal{L} \equiv \lambda/N(\Delta p_i) = \lambda \mathcal{T}(0) = \lambda \alpha_0$. Thus

$$\ell = \left(\sum_{i=1}^{\infty} \frac{\Gamma\left(\frac{n}{2}\right)}{i!\,\Gamma\left(\frac{n}{2}+i\right)} \left(\frac{\lambda \bar{y}}{2\,\sigma^2}\right)^i\right) \Big/ \left(\sum_{i=0}^{\infty} \frac{\Gamma\left(\frac{n}{2}\right)}{i!\,\Gamma\left(\frac{n}{2}+i\right)} \left(\frac{\lambda \bar{y}}{2\,\sigma^2}\right)^i\right)$$

which is equation (6). □

Proof. (of Proposition 4). We first state a lemma about the density $f(y)$.

Lemma 3 *Let $f(y; n, \frac{\lambda}{\sigma^2}, \bar{y})$ be the density of $y \in [0, \bar{y}]$ in equation (7) satisfying the boundary conditions. For any $k > 0$ we have: $f\left(y; n, \frac{\lambda}{\sigma^2}, \bar{y}\right) = \frac{1}{k}\ f\left(\frac{y}{k}; n, \frac{\lambda k}{\sigma^2}, \frac{\bar{y}}{k}\right)$.*

Proof. (of Lemma 3). Consider the function $f(y; n, \frac{\lambda}{\sigma^2}, \bar{y})$ solving equation (7) (and boundary conditions) for given $n, \frac{\lambda}{\sigma^2}, \bar{y}$. Without loss of generality set $\sigma' = \sigma$ and consider $\bar{y}' = \bar{y}/k$ and $\lambda' = \lambda k$. Notice that by setting $C_1' = C_1 k$ and $C_2' = C_2 k$ we verify that the boundary conditions hold (because $C_1'/C_2' = C_1/C_2$) and that (7) holds (which is readily verified by a change of variable). □

We now prove the proposition. Let $w\left(\Delta p_i; n, \ell, Std(\Delta p_i)\right)$ be the density function in equation (9). Next we verify equation (10). From the first term in equation (9) notice that

$$(1-\ell)\,\omega\left(\Delta p_i; \bar{y}\right) = s\ (1-\ell)\,\omega\left(s\Delta p_i; s^2 \bar{y}\right)$$

where the first equality uses the homogeneity of degree -1 of $\omega(\Delta p_i; y)$ (see equation (8)). From the second term in equation (9) for $n \geq 2$

$$\ell \int_0^{\bar{y}} \omega(\Delta p_i; y) f(y; n, \frac{\lambda}{\sigma^2}, \bar{y}) dy = \ell \int_0^{\bar{y}} s\ \omega\left(s\Delta p_i; s^2 y\right)\ \ s^2\ f\left(y s^2; n, \frac{\lambda}{s^2 \sigma^2}, \bar{y} s^2\right) dy$$

where the first equality follows from Lemma 3 for $k = 1/s^2$, and the homogeneity of degree -1 of $\omega(\cdot, \cdot)$. Further we note

$$\ell \int_0^{\bar{y}} s\ \omega\left(s\Delta p_i; s^2 y\right) s^2\ f\left(y s^2; n, \frac{\lambda}{s^2 \sigma^2}, \bar{y} s^2\right) dy = s^3 \ell \int_0^{\bar{y}} \omega\left(s\Delta p_i; s^2 y\right)\ f\left(y s^2; n, \frac{\lambda'}{\sigma'^2}, \bar{y}'\right) dy$$

where $\frac{\lambda' \bar{y}'}{\sigma'^2} = \frac{\lambda \bar{y}}{\sigma^2}$, so that ℓ is the same across the two economies. Using $z = y\ s^2$

$$s^3 \ell \int_0^{\bar{y}} \omega\left(s\Delta p_i; s^2 y\right) f\left(ys^2; n, \frac{\lambda'}{\sigma'^2}, \bar{y}'\right) dy = s\ \ell \int_0^{\bar{y}'} \omega\left(s\Delta p_i; z\right) f\left(z; n, \frac{\lambda'}{\sigma'^2}, \bar{y}'\right) dz\ .$$

where $\bar{y}' = s^2 \bar{y}$, which completes the verification of equation (10). $\square$

Proof. (of Proposition 5). For any $p \in \mathbb{R}^n$ with $||p||^2 \leq \bar{y}$, we write $m(p; \bar{y}, \sigma, \lambda)$ to emphasize the dependence on $(\bar{y}, \sigma, \lambda)$. A guess and verify strategy can be used to show the following scaling property of the function m: Let $k > 0$, then for all $p \in \mathbb{R}^n$ with $||p||^2 \leq \bar{y}$:

$$m(kp; k^2\bar{y}, k\sigma, \lambda) = k\, m(p; \bar{y}, \sigma, \lambda) \quad \text{and} \quad m(p; \bar{y}, \sigma\sqrt{k}, \lambda k) = \frac{1}{k}\, m(p; \bar{y}, \sigma, \lambda)\,.$$

It is straightforward to verify that this function satisfies the ODE and boundary conditions for $m(p)$ (see e.g. the one in the main text for the $n = 1$ case). Recall the homogeneity of $f(y)$ stated in Lemma 3. Finally, note that the density $g(p)$ can be expressed as a function of the density $f(y)$ given in equation (7) and the density of the sum of n coordinates of a random variable uniformly distributed on a n dimensional hypersphere of square radius y, as obtained in Equation 21 in Alvarez and Lippi (2014). These properties applied to equation (16) establish the scaling property stated in the proposition.

Proof. (of Proposition 6). We first notice that for some special cases a simple analytic proof is available. These cases concern $n = 1$ or $n = \infty$ with $\ell \in (0, 1)$; alternatively, they concern $1 < n < \infty$ and $\ell = 0$ or $\ell = 1$. See the technical Appendix G for details.

We now assume $1 \leq n < \infty$ and $0 < \ell < 1$ and prove that $\mathcal{M}'(0) = \frac{Kur(\Delta p_i)}{6\, N(\Delta p_i)}$. The proof is structured as follows. First we derive an analytic expressions for $\frac{Kur(\Delta p_i)}{6\, N(\Delta p_i)}$ and for $\mathcal{M}'(0)$. Each expression is a power series that involves only two parameters: n and ℓ. Verifying the equality is readily done numerically to any arbitrary degree of accuracy. A simple Matlab code for the verification, called `solveMp0.m`, is available on our websites.

We first show that $\frac{Kur(\Delta p_i)}{6\, N(\Delta p_i)}$ can be written as the right hand side of equation (19). This is done in two steps. First notice that $\frac{Kur(\Delta p_i)}{N(\Delta p_i)}$ $=(1/\lambda)\mathcal{L}(\phi, n)\, Kur\left(\Delta p_i\right)$ where $\mathcal{L}$ is given in Proposition 3. The second step is to derive an analytic expression for the $Kur\left(\Delta p_i\right)$. We notice that

$$Kur\left(\Delta p_i\right) = \frac{\mathbb{E}\left(\Delta p_i^4(\tau) \,|\, y(0) = 0\right)}{Var\left(\Delta p_i\right)^2} = \frac{Q(0)}{\frac{\sigma^4}{N(\Delta p_i)^2}} = \frac{\left(\lambda/\sigma^2\right)^2 Q(0)}{\left(\mathcal{L}(\phi, n)\right)^2}$$

where τ is the stopping time associated with a price change, and where $Q(y)$ is the expected fourth moment at the time of adjustment τ conditional on a current squared price gap y, i.e.

$$Q(y) = \mathbb{E}\left(\Delta p_i^4(\tau)\,|\,y(0)=y\right) = \frac{3}{(n+2)n}\,\mathbb{E}\left(y^2(\tau)|y(0)=y\right)$$

where $y(\tau)$ is the value of the squared price gap at the stopping time. Notice that for $y \in [0,\bar{y}]$ the function $Q(y)$ obeys the o.d.e.:

$$\lambda Q(y) = \lambda \frac{3y^2}{(n+2)n} + Q'(y)\,n\sigma^2 + Q''(y)\,2\sigma^2 y$$

with boundary condition $Q(\bar{y}) = \frac{3\bar{y}^2}{(n+2)n}$. The solution of Q has a power series representation which is easily obtained by matching coefficients and using the boundary conditions. Using this power series in the expression for $\frac{Kur(\Delta p_i)}{N(\Delta p_i)}$ obtained in the first step gives the expression on the right hand side of equation (19). See the technical Appendix L for details on the algebra.

Next we derive an expression for $\mathcal{M}'(0)$, which holds for all $1 \le n < \infty$ and $0 \le \ell \le 1$:

Lemma 4 *Let $\mathcal{M}(\cdot;n)$ be the area under the IRF of output and $f(\cdot;n)$ be the density of the invariant distribution for an economy with n products and parameters $(\bar{y},\lambda,\sigma^2)$. Let $\mathcal{T}_{n+2}(y)$ be the expected time until either $y(t)$ hits $\bar{y}$ or that until there is a free adjustment opportunity, whichever happens first, starting at $y(0)=y$, for an economy with $n+2$ products and the same parameters $(\bar{y},\lambda,\sigma^2)$. Then*

$$\mathcal{M}'(0;n) = \frac{1}{\epsilon}\int_0^{\bar{y}} \left[\mathcal{T}_{n+2}(y) + \frac{2}{n}\,\mathcal{T}'_{n+2}(y)\,y\,\right] f(y;n)\;dy \quad . \tag{26}$$

The function $\mathcal{T}_n(y)$ is characterized in the proof of Proposition 3 where we give an explicit power series representation for this function. The proof of Lemma 4 uses a characterization of $m(p)$ in terms of a two dimensional vector (z,y), where z is the sum of the n coordinates of p. The function $m(z,y)$ solves a PDE whose solution can be expressed in terms of $\mathcal{T}_n(y)$ (see the technical Appendix M for details on the algebra).

To compute the right hand side of equation (26), we separately characterize $\mathcal{T}_{n+2}(y) + \frac{2}{n}\,\mathcal{T}'_{n+2}(y)\,y$ and $f(y;n)$. Using the power series representation of $\mathcal{T}_{n+2}(y)$ (see proof of Proposition 3) it is immediate to obtain a power series representation of $\mathcal{T}_{n+2}(y) + \frac{2}{n}\,\mathcal{T}'_{n+2}(y)\,y$. This gives (see the technical Appendix N for details on the algebra):

$$\mathcal{T}_{n+2}(y) + \frac{2}{n}\,\mathcal{T}'_{n+2}(y)\,y = \frac{\sum_{i=1}^{\infty} \frac{\Gamma\left(\frac{n}{2}+1\right)}{i!\,\Gamma\left(\frac{n}{2}+i+1\right)} \left[\left(\frac{\lambda\bar{y}}{2\,\sigma^2}\right)^i - \left(1+\frac{2i}{n}\right)\,\left(\frac{\lambda y}{2\sigma^2}\right)^i\right]}{\sum_{i=0}^{\infty} \frac{\Gamma\left(\frac{n}{2}+1\right)}{i!\,\Gamma\left(\frac{n}{2}+i+1\right)}\left(\frac{\lambda\bar{y}}{2\,\sigma^2}\right)^i} \tag{27}$$

For f we use the characterization in equation (7) in term of modified Bessel functions of the first and second kind (see the technical Appendix O for a step-by-step derivation). These functions have a power series representation, which we use to solve for the two unknown constants C_1, C_2. This gives:

$$f(y) \quad = \quad \left[\frac{\left(\frac{\lambda y}{2\sigma^2}\right)^{\left(\frac{n}{2}-1\right)} \; \sum_{i=0}^{\infty} \beta_{i,\frac{n}{2}-1} \left(\frac{\lambda y}{2\sigma^2}\right)^i}{\left(\frac{\lambda \bar{y}}{2\sigma^2}\right)^{\left(\frac{n}{2}-1\right)} \; \sum_{i=0}^{\infty} \beta_{i,\frac{n}{2}-1} \left(\frac{\lambda \bar{y}}{2\sigma^2}\right)^i} - \frac{\sum_{i=0}^{\infty} \beta_{i,1-\frac{n}{2}} \left(\frac{\lambda y}{2\sigma^2}\right)^i}{\sum_{i=0}^{\infty} \beta_{i,1-\frac{n}{2}} \left(\frac{\lambda \bar{y}}{2\sigma^2}\right)^i} \right] \Bigg/ \tag{28}$$

$$\left[\frac{\left(\frac{\lambda \bar{y}}{2\sigma^2}\right)^{\left(\frac{n}{2}-1\right)} \; \sum_{i=0}^{\infty} \beta_{i,\frac{n}{2}-1} \frac{\bar{y}}{i+n/2} \left(\frac{\lambda \bar{y}}{2\sigma^2}\right)^i}{\left(\frac{\lambda \bar{y}}{2\sigma^2}\right)^{\left(\frac{n}{2}-1\right)} \; \sum_{i=0}^{\infty} \beta_{i,\frac{n}{2}-1} \left(\frac{\lambda \bar{y}}{2\sigma^2}\right)^i} - \frac{\sum_{i=0}^{\infty} \beta_{i,1-\frac{n}{2}} \frac{\bar{y}}{i+1} \left(\frac{\lambda \bar{y}}{2\sigma^2}\right)^i}{\sum_{i=0}^{\infty} \beta_{i,1-\frac{n}{2}} \left(\frac{\lambda \bar{y}}{2\sigma^2}\right)^i} \right]$$

where the two sequence of coefficients β are defined in term of the Γ function as

$$\beta_{i,\frac{n}{2}-1} \equiv \frac{1}{i! \; \Gamma(i+n/2)} \quad \text{and} \quad \beta_{i,1-\frac{n}{2}} \equiv \frac{1}{i! \; \Gamma(i+2-n/2)} \quad \text{for } i = 0, 1, 2, \ldots \quad .$$

The expression in equation (28) holds for all *real* numbers $n \geq 1$, except when n is an even natural number (due to a singularity of the power expansion of the modified Bessel function of the second kind). Yet the expression is continuous in n.

Finally, we establish an equivalence to verify equation (19):

Lemma 5 *The equality between equation (26) and the ratio $Kur(\Delta p_i)/(6N(\Delta p_i))$, as from equation (19), is equivalent to*

$$\sum_{j=1}^{\infty} \frac{\gamma_j \, \frac{1}{1+j}}{\sum_{s=0}^{\infty} \gamma_s \, \frac{1}{1+s}} \, j \quad = \quad \sum_{j=1}^{\infty} \gamma_j \left(1 + \frac{2\,j}{n}\right) \times \tag{29}$$

$$\left(\left[\frac{\sum_{i=0}^{\infty} \xi_i \, \frac{1}{\frac{n}{2}+i+j}}{\sum_{i=0}^{\infty} \xi_i} - \frac{\sum_{i=0}^{\infty} \rho_i \, \frac{1}{i+1+j}}{\sum_{i=0}^{\infty} \rho_i} \right] \Bigg/ \left[\frac{\sum_{i=0}^{\infty} \xi_i \, \frac{1}{\frac{n}{2}+i}}{\sum_{i=0}^{\infty} \xi_i} - \frac{\sum_{i=0}^{\infty} \rho_i \, \frac{1}{i+1}}{\sum_{i=0}^{\infty} \rho_i} \right] \right),$$

where the sequences $\{\gamma_j, \xi_j, \rho_j\}_{j=0}^{\infty}$ are defined as

$$\gamma_j \quad \equiv \quad \frac{\Gamma\left(\frac{n}{2}+1\right)}{j! \, \Gamma\left(\frac{n}{2}+1+j\right)} \left(\frac{\lambda \bar{y}}{2\sigma^2}\right)^j, \; \xi_j \equiv \frac{1}{j! \, \Gamma\left(j+\frac{n}{2}\right)} \left(\frac{\lambda \bar{y}}{2\sigma^2}\right)^{\left(\frac{n}{2}+j-1\right)} \quad \text{and} \quad \rho_j \equiv \frac{1}{j! \, \Gamma\left(j+2-\frac{n}{2}\right)} \left(\frac{\lambda \bar{y}}{2\sigma^2}\right)^j.$$

The derivation of equation (29) uses equation (27) and equation (28) to compute equation (26). Verifying equation (29) is straightforward since both sides are simple functions of convergent power series, which are arbitrarily well approximated by a finite sum. As explained above, for even values of n this expression should be understood as the limit for $n \to 2\,k$ (or, numerically, as the sum for values of n close to $2\,k$ for $k \in \mathbb{N}$ and $k \geq 1$). $\square$

Proof. (of Proposition 7). The idea is to show that for any n and ℓ we have

$$Kur(\Delta p_i;\mu) = Kur(\Delta p_i;-\mu)\ ,\ N(\Delta p_i)(\mu) = N(\Delta p_i)(-\mu)\ ,\ \mathcal{M}(\delta;\mu) = -\mathcal{M}(-\delta;-\mu) \tag{30}$$

for all (μ,δ) in a neighborhood of $(0,0)$. Note that differentiating the last expression with respect to δ, and evaluating it at $\delta = 0$ we obtain that $\mathcal{M}'(0;\mu) = \mathcal{M}'(0;-\mu)$. Hence we have that $Kur(\Delta p_i;\cdot), N(\Delta p_i)(\cdot)$ and $\mathcal{M}'(0;\cdot)$ are symmetric functions of inflation around $\mu = 0$. Hence, if they are differentiable, they must have zero derivative with respect to inflation at zero inflation. The symmetry in equation (30) follows from the symmetry on the firm's problem with respect to positive and negative drift. To establish this symmetry we proceed in two steps. First we analyze the symmetry of the decision problem for the firm of Section 3.2. Second, we consider the approximation to the GE problem for values $\mu \neq 0$. All the arguments follow a guess and verify strategy of a simple nature but with heavy notation. The technical Appendix S provides the details of the proof.

Proof. (of Proposition 8.) The proof proceeds by verification. We analyze the condition that ensures that every firm with $||p||^2 = y \leq \bar{y}$ before the shock will find that $||p - \iota\delta||^2 \geq \bar{y}$ after the shock, where ι is a vector of ones. See the technical Appendix Q for a detailed derivation.

Proof. (of Lemma 2) To prove the lemma we let $z(t) = x^4(t)$ and $y(t) = x^2(t)$. We use Ito's lemma to obtain:

$$dy(t) = \sigma(t)^2\,dt \ + \ 2x(t)\,\sigma(t)dW(t) \qquad \text{or} \quad y(t) = \int_0^t \sigma(s)^2\,ds \ + \ \int_0^t 2\,x(s)\,\sigma(s)\,dW(s)$$

and taking expected values: $\mathbb{E}_0\left[y(T)\right] = \mathbb{E}_0\left[\int_0^T \sigma(t)^2\,dt\right] = \int_0^T \mathbb{E}_0\left[\sigma(t)^2\right]dt$. Likewise

$$dz(t) = 6\,x^2(t)\,\sigma(t)^2\,dt \ + \ 4x^3(t)\,\sigma(t)dW(t)$$

then $\mathbb{E}_0\left[z(T)\right] = 6\,\mathbb{E}_0\left[\int_0^T \sigma(t)^2\,y(t)\,dt\right] = 6\,\int_0^T \mathbb{E}_0\left[\sigma(t)^2\,y(t)\right]dt$. Now note that:

$$\mathbb{E}_0\left[\sigma(t)^2\,y(t)\,dt\right] = \mathbb{E}_0\left[\sigma(t)^2\left(\int_0^t \sigma(s)^2\,ds \ + \ \int_0^t 2\,x(s)\,\sigma(s)\,dW(s)\right)\right]$$

and using the independence of $\{W(t)\}$ and $\{\sigma(t)\}$ we have: $\mathbb{E}_0\left[\sigma(t)^2\,y(t)\right] = \mathbb{E}_0\left[\int_0^t \sigma(t)^2\,\sigma(s)^2\,ds\right]$. Then, replacing this expression and noticing that $K(T) = \mathbb{E}_0\left[z(T)\right]/\left(\mathbb{E}_0\left[y(T)\right]\right)^2$, we obtain the desired result. □

Proof. (of Proposition 9.) Define the probability that $u(t) = 1$ if $u(0) = i$ as $P_1\left(t\,|\,i\right)$, or:

$P_1\left(t\,|\,i\right) \equiv \Pr\left\{u(t)=1 \mid u(0)=i\right\}$ for $i \in \{0,1\}$. These probabilities are given by:

$$P_1\left(t\,|\,0\right) = \frac{\theta_0}{\theta_0+\theta_1}\left[1-e^{-(\theta_0+\theta_1)t}\right] \quad , \quad P_1\left(t\,|\,1\right) = \frac{\theta_0}{\theta_0+\theta_1}\left[1+\frac{\theta_1}{\theta_0}e^{-(\theta_0+\theta_1)t}\right] .$$

We now use these probabilities to compute two expressions that appear in equation (21). (see the technical Appendix T for the details on the derivation). The first expression is the expected second moment $v(t) \equiv \mathbb{E}_0\left[\sigma(t)^2\right]$ which is given by

$$v(t) = \mathbb{E}_0\left[\sigma(t)^2\right] = \sigma_0^2\frac{\theta_1}{\theta_1+\theta_0} + \sigma_1^2\frac{\theta_0}{\theta_0+\theta_1} . \tag{31}$$

Notice that this expected variance is independent of the horizon t. The second expression is $k(t,s) \equiv \mathbb{E}_0\left[\sigma(t)^2\,\sigma(s)^2\right]$, or the expected fourth moment over an horizon t conditional on $\sigma(t)$. This is given by

$$k(t,s) = \left[\sigma_0^2\frac{\theta_1}{\theta_1+\theta_0} + \sigma_1^2\frac{\theta_0}{\theta_1+\theta_0}\right]^2 + \left(\sigma_1^2-\sigma_0^2\right)^2\,\frac{\theta_0\theta_1}{(\theta_0+\theta_1)^2}\,e^{-(\theta_0+\theta_1)(t-s)} \tag{32}$$

Using equation (31) and equation (32) into equation (21) gives

$$K(T) = 3 + 6\,\frac{\left(\sigma_1^2-\sigma_0^2\right)^2\,\frac{\theta_0\theta_1}{(\theta_0+\theta_1)^2}}{\left[\sigma_0^2\frac{\theta_1}{\theta_1+\theta_0}+\sigma_1^2\frac{\theta_0}{\theta_1+\theta_0}\right]^2}\,\frac{\left[T\left(\theta_0+\theta_1\right)-1+e^{-(\theta_1+\theta_0)T}\right]}{(\theta_1+\theta_0)^2\,T^2} .$$

Without loss of generality, since the expression is homogeneous of degree zero on (σ_1,σ_0), we can set $\sigma_1 = 1$. We can also use $\theta = (1/2)(\theta_1+\theta_0)$. Finally for any θ we can let $s = \theta_0/(\theta_0+\theta_1)$ to obtain equation (22). $\square$

Proof. (of Proposition 10.) First we compute the frequency of adjustment. Let $T_i(p)$ denote the expected time to hit a barrier conditional on the state p. The Kolmogorov backward equation gives the following system of ODEs for the expected times:

$$\begin{cases} \theta_0(T_0-T_1) = 1 + T_0''\,\frac{\sigma_0^2}{2} \\ \theta_1(T_1-T_0) = 1 + T_1''\,\frac{\sigma_1^2}{2} \end{cases}$$

which is symmetric $T_i(p) = T_i(-p)$ with boundary condition $T_i(\bar{p}) = 0$. The solution is

$$\begin{cases} T_0(p) = \frac{(\theta_0+\theta_1)(\bar{p}^2-p^2)}{\sigma_0^2\theta_1+\sigma_1^2\theta_0} + \frac{\sigma_1^2\theta_0(\sigma_0^2-\sigma_1^2)}{(\sigma_0^2\theta_1+\sigma_1^2\theta_0)^2}\,\left(\frac{e^{\chi p}+e^{-\chi p}}{e^{\chi\bar{p}}+e^{-\chi\bar{p}}}-1\right) \\ T_1(p) = \frac{(\theta_0+\theta_1)(\bar{p}^2-p^2)}{\sigma_0^2\theta_1+\sigma_1^2\theta_0} + \frac{\sigma_0^2\theta_1(\sigma_1^2-\sigma_0^2)}{(\sigma_0^2\theta_1+\sigma_1^2\theta_0)^2}\,\left(\frac{e^{\chi p}+e^{-\chi p}}{e^{\chi\bar{p}}+e^{-\chi\bar{p}}}-1\right) \end{cases} \qquad \text{where} \;\; \chi \equiv \sqrt{2\frac{\sigma_0^2\theta_1+\sigma_1^2\theta_0}{\sigma_0^2\sigma_1^2}}$$

This implies that the average time between price adjustment is given by

$$\frac{1}{N_a} = \frac{T_0(0)\theta_1 + T_1(0)\theta_0}{\theta_0 + \theta_1} = \frac{(\theta_0+\theta_1)\bar{p}^2}{\sigma_0^2\theta_1 + \sigma_1^2\theta_0} + \frac{\theta_0\theta_1(\sigma_1^2-\sigma_0^2)^2}{(\theta_0+\theta_1)(\sigma_0^2\theta_1+\sigma_1^2\theta_0)^2}\left(1 - \frac{2}{e^{\chi\bar{p}} + e^{-\chi\bar{p}}}\right)$$

or, rewriting in terms of the fundamental parameters that pin down $K(T)$, namely θ, s, ξ, and the implied parameter $\hat{\rho} = \frac{\theta_1}{\theta_0} = \frac{1-s}{s}$ we have equation (23).

Now we turn to computing the cumulative output effect. Use the approximation

$$\mathcal{M}(\delta) \quad \approx \quad \delta\ \mathcal{M}'(0) = \frac{2\delta}{\epsilon}\int_0^{\bar{p}} \left(m_0(p)g_0'(p) + m_1(p)g_1'(p)\right)\, dp$$

Next we solve for the terms in the equation. First consider the ODE that characterizes $m_i(p)$:

$$\begin{cases} \theta_0(m_0 - m_1) = -p + \frac{\sigma_0^2}{2}m_0'' \\ \theta_1(m_1 - m_0) = -p + \frac{\sigma_1^2}{2}m_1'' \end{cases}$$

The function must satisfy $m_i(p) = -m_i(-p)$ and the boundaries $m_i(\bar{p}) = 0$. The solution is

$$\begin{cases} m_0(p) = \frac{(\theta_0+\theta_1)\,p\left(p^2-\bar{p}^2\right)}{3(\sigma_0^2\theta_1+\sigma_1^2\theta_0)} + \frac{\sigma_1^2\theta_0(\sigma_1^2-\sigma_0^2)}{(\sigma_0^2\theta_1+\sigma_1^2\theta_0)^2}\left(\frac{e^{\chi p}-e^{-\chi p}}{e^{\chi\bar{p}}-e^{-\chi\bar{p}}}\,\bar{p} - p\right) \\ m_1(p) = \frac{(\theta_0+\theta_1)\,p\left(p^2-\bar{p}^2\right)}{3(\sigma_0^2\theta_1+\sigma_1^2\theta_0)} + \frac{\sigma_0^2\theta_1(\sigma_0^2-\sigma_1^2)}{(\sigma_0^2\theta_1+\sigma_1^2\theta_0)^2}\left(\frac{e^{\chi p}-e^{-\chi p}}{e^{\chi\bar{p}}-e^{-\chi\bar{p}}}\,\bar{p} - p\right) \end{cases} \quad \text{where} \quad \chi \equiv \sqrt{2\frac{\sigma_0^2\theta_1+\sigma_1^2\theta_0}{\sigma_0^2\sigma_1^2}}$$

Finally we compute the invariant distribution of price gaps. Let $g_i(p)$ be the density for price gaps in state i which must be symmetric around $p = 0$, zero at the boundary: $g_i(\bar{p}) = 0$.

$$\begin{cases} \frac{\sigma_0^2}{2}g_0''(p) = \theta_1 g_1(p) - \theta_0 g_0(p) \\ \frac{\sigma_1^2}{2}g_1''(p) = \theta_0 g_0(p) - \theta_1 g_1(p) \end{cases}$$

For $p \in [-\bar{p}, \bar{p}]$, the shape of the densities is linear triangular, with density functions

$$\begin{cases} g_0(p) = \frac{\theta_1}{\theta_0+\theta_1}\frac{\bar{p}-|p|}{\bar{p}^2} \\ g_1(p) = \frac{\theta_0}{\theta_0+\theta_1}\frac{\bar{p}-|p|}{\bar{p}^2} \end{cases}$$

□

C On the implied cost of price adjustment

In this section we give a characterization of the model implications for the size of the menu cost, i.e. a mapping between observable statistics and the value of ψ/B or ψ (we also discuss how to measure B). We consider two measures for the cost of price adjustment: the first one is the cost of a single price adjustment as a fraction of profits: ψ/n. Recall that ψ is the cost that a firm must pay if it decides to adjust all prices instantaneously (i.e. without waiting for a free adjustment). Measuring this cost as a fraction of profits transforms these magnitudes into units that have an intuitive interpretation. The second measure is the average flow cost of price adjustment given by: $N(\Delta p_i)\ \frac{\psi}{n}\,(1-\ell)$. This cost measures the average amount of resources that the firm pays to adjust prices per period. The latter measure is useful because it relates more directly to what has been measured in the data by Levy et al. (1997); Zbaracki et al. (2004), namely the "average" cost of a price adjustment. The next proposition analyzes the mapping between the scaled menu cost ψ/n , and B, ℓ, n, $N(\Delta p_i)$ and $Var(\Delta p_i)$.

Proposition 11 *Fix the number of products $n \geq 1$ and let $r \downarrow 0$. There is a unique triplet $(\sigma^2, \lambda, \psi)$ consistent with any triplet $\ell \in [0,1], Var(\Delta p_i) > 0$ and $N(\Delta p_i) > 0$. Moreover, fixing any value ℓ, the menu cost $\psi \geq 0$ can be written as:*

$$\frac{\psi}{n} = B\ \frac{Var(\Delta p_i)}{N(\Delta p_i)}\ \Psi\,(n\,,\,\ell) \tag{33}$$

where Ψ is only a function of (n, ℓ). For all $n \geq 1$ the function $\Psi(n, \cdot)$ satisfies:

$$\lim_{\ell\to 0} \Psi\,(n,\,\ell) = \frac{n}{2\,(n+2)}\ ,\quad \lim_{\ell\to 1} \Psi\,(n,\,\ell) = \infty\ ,\quad \lim_{\ell\to 1} \Psi\,(n,\,\ell)\,(1-\ell) = 0\ , \tag{34}$$

$$\lim_{\ell\to 1} \frac{\Psi(n', \ell)/n'}{\Psi(n,\ell)/n} \leq 1 \ \text{ for } n' \geq n, \ \text{ and } \lim_{n\to\infty} \frac{\Psi(n,\ell)/n}{\Psi(1,\ell)/1} \to 0 \text{ as } \ell \to 1\,. \tag{35}$$

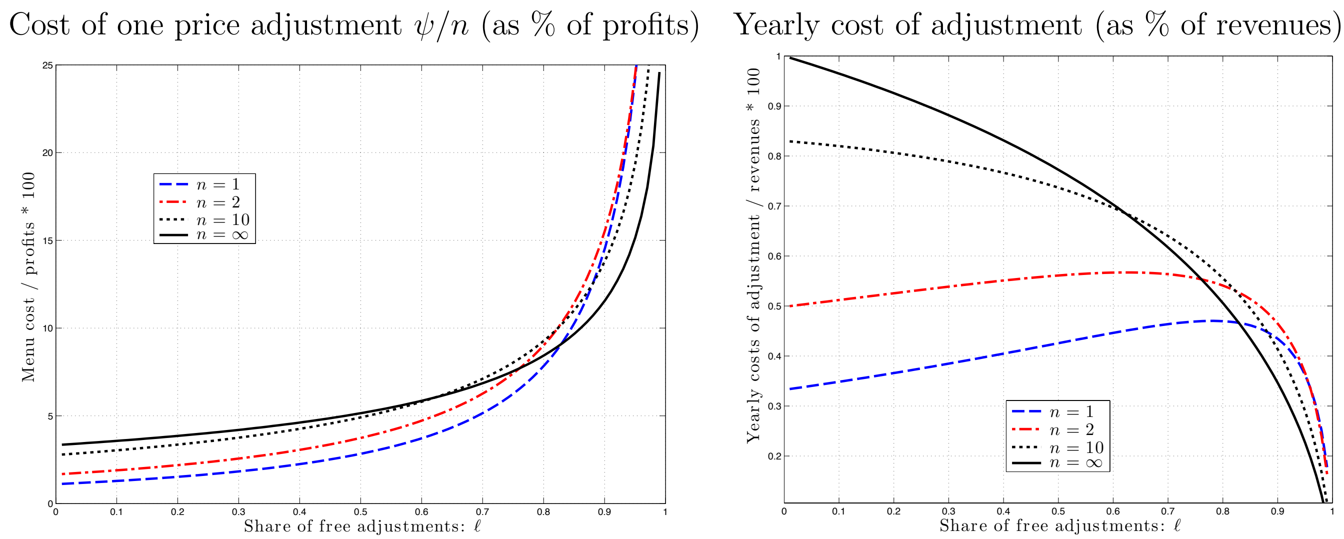
Equation (33) shows that for any fixed $n \geq 1$ and $\ell \in [0,1]$ the menu cost ψ is proportional to the ratio $Var(\Delta p_i)/N(\Delta p_i)$. Second, equation (33) shows that the menu cost is proportional to B, which measures the benefits of closing a price gap. The parameter B is related to the constant demand elasticity faced by firms η (see Section 3), so that $B = \eta(\eta-1)/2$, which can be written in terms of the (net) markup over marginal costs $\mathfrak{m} \equiv 1/(\eta-1)$ so that $B = (1+\mathfrak{m})/(2\mathfrak{m}^2)$.[28] The last expression is useful to calibrate the model using empirical estimates of the markup such as the ones by Christopoulou and Vermeulen (2012): the estimated markups average around 28% for the US manufacturing sector, and around 36% for market services (slightly smaller values are obtained for France, see their Table 1).[29] A similar value for the US, namely a markup rate of about 33%, is used by Nakamura and

[28] Nakamura and Steinsson (2010) notice that lower markups (higher values of demand elasticity) η must imply higher menu costs, as shown by equation (33). Footnote 14 in their paper discusses evidence on the markup rates across several microeconomic studies and macro papers.

[29] The evidence for the US services is consistent with the gross margins, based on accounting data, reported in the Annual Retail Trade Survey by the US Census (see http://www.census.gov/retail/).

Steinsson (2010).

Figure 6: Implied cost of price adjustment



All economies in the figures feature $Std(\Delta p_i) = 0.10$ and a markup of 25%. For those in the left panel we set $N(\Delta p_i) = 1.5$.

The left panel of Figure 6 illustrates the comparative static effect of ℓ and n on the implied menu cost, fixing $B\,Var(\Delta p_i)/\,N(\Delta p_i)$, i.e. it plots the function $\Psi(n, \ell)$. Fixing a value of n it can be seen that the menu cost ψ/n is increasing in ℓ. Indeed equation (34) shows that as $\ell \to 1$, the implied menu cost diverges to $+\infty$. On the other hand, for $\ell = 0$ and $n = 1$, our version of *Golosov-Lucas*'s model, the menu cost attains its smallest (strictly positive) value. Fixing ℓ and moving across lines shows that the implied fixed cost ψ/n is not monotone in the number of products n. Indeed, as stated in equation (34) for a very small share ℓ the values of ψ/n are increasing in n. On the other hand, for larger value of the share ℓ, the order of the implied fixed cost is reversed.

The model also has clear predictions about the per period (say yearly) cost of price adjustments borne by the firms: $(1-\ell)\,N(\Delta p_i)\,\psi/n$. In spite of the fact that the cost of a single deliberate price adjustment diverges as $\ell \to 1$, the total yearly cost of adjustment converge to zero continuously. This can be seen in the right panel of Figure 6. A simple transformation gives the yearly cost of price adjustments as a fraction of revenues: $\frac{(1-\ell)N(\Delta p_i)\ \psi/n}{\eta}$, where the scaling by η transforms the units from fraction of profits into fraction of revenues.[30] This statistic is useful because it has empirical counterparts, studied e.g. by Levy et al. (1997).

[30] Since $R = \eta\,\Pi$ where R is revenues per good and Π profits per good.

Using equation (33) and the previous definition for the markup yields

$$\frac{\text{Yearly costs of price adjustment}}{\text{Yearly revenues}} = \frac{1}{2} \frac{Var(\Delta p_i)}{\mathfrak{m}} \, (1-\ell) \, \Psi \, (n \, , \, \ell) \tag{36}$$

Figure 6 plots the two cost measures in equation (33) and (36) as functions of ℓ, n for an economy with $N(\Delta p_i) = 1.5$, $Std(\Delta p_i) = 0.10$ and a markup $\mathfrak{m} \approx 25\%$ (i.e. $B = 10$). We see this parametrization as being consistent with the US data on price adjustments, markups, and the size distribution of price changes discussed above. The figure illustrates how observations on the costs of price adjustments can be used to parametrize the model. Levy et al. (1997) and Dutta et al. (1999) (Table IV and Table 3, respectively) document that for multi-product stores (a handful of supermarket chains and one drugstore chain) the average cost of price adjustment is around 0.7 percent of revenues. For an economy with $n = 10$ (a reasonable parametrization to fit the size-distribution of price changes) the right panel of the figure shows that the model reproduces the yearly cost of 0.7% of revenues when the fraction of free adjustments ℓ is around 60%. The left panel in the figure indicates that at this level of ℓ the cost of one price adjustment is around 5% of profits.

Proof. (of Proposition 11). To obtain the expression in equation (33) we use the characterization of $\ell = \mathcal{L}\left(\frac{\lambda \bar{y}}{n\sigma^2}, n\right)$ of Proposition 3, it is equivalent to fix a value of $\phi \equiv \frac{\lambda \bar{y}}{n\sigma^2}$. We let the optimal decision rule be $\bar{y}\left(\psi/B, \sigma^2, r+\lambda, n\right)$ so that we have:

$$\bar{y}\left(\frac{\psi}{B}, \sigma^2, r+\lambda, n\right) \, \frac{\lambda}{n\,\sigma^2} = \phi$$

To be consistent with $Var(\Delta p_i)$ and $N(\Delta p_i)$ we have, using Proposition 1 and $\ell = \mathcal{L}(\phi, n)$:

$$N(\Delta p_i) = \lambda / \mathcal{L}(\phi, n) \;\; \text{and} \;\; \frac{\lambda}{\sigma^2} = \mathcal{L}(\phi, n)/Var(\Delta p_i) \; .$$

Thus, after taking $r \downarrow 0$ and using the expression above we can write:

$$\bar{y}\left(\frac{\psi}{B} \, , \, N(\Delta p_i) Var(\Delta p_i) \, , \, \ell \, N(\Delta p_i) \, , \, n\right) \, \frac{\ell}{n \, Var(\Delta p_i)} = \mathcal{L}^{-1}(\ell; n)$$

Fixing n and ℓ and computing the total differential for this expression with respect to $(\psi/B, N(\Delta p_i), Var(\Delta p_i))$, and denoting by η_ψ, η_{σ^2}, η_λ the elasticities of $\bar{y}$ with respect to ψ/B, σ^2, λ we have:

$$\eta_\psi \, \hat{\psi} + \eta_{\sigma^2} \, (\hat{N}(\Delta p_i) + \hat{V}ar(\Delta p_i)) + \eta_\lambda \, \hat{N}(\Delta p_i) = \hat{V}ar(\Delta p_i)$$

where a *hat* denotes a proportional change. Using Proposition 3-(iv) in Alvarez and Lippi (2014) and Lemma 1 we have that these elasticities are related by: $\eta_\lambda = 2\eta_\psi - 1$ and $\eta_{\sigma^2} = 1 - \eta_\psi$ .Thus $\eta_\psi \, \hat{\psi} + (1-\eta_\psi) \, (\hat{N}(\Delta p_i) + \hat{V}ar(\Delta p_i)) + (2\eta_\psi - 1) \, \hat{N}(\Delta p_i) = \hat{V}ar(\Delta p_i)$. Rearranging and canceling terms: $\eta_\psi \, \hat{\psi} + \eta_\psi \, \hat{N}(\Delta p_i) - \eta_\psi \hat{V}ar(\Delta p_i) = 0$. Dividing by η_ψ we obtain that $\hat{\psi} = \hat{V}ar(\Delta p_i) - \hat{N}(\Delta p_i)$. Additionally, since $\bar{y}$ is a function of ψ/B, then we can write $\psi/n = B \, (Var(\Delta p_i)/N(\Delta p_i)) \; \Psi(n, \ell)$ for some function $\Psi(n, \ell)$.

That $\psi \to \infty$ as $\ell \to 1$ follows because $\mathcal{L}(\phi, n) \to 1$ as $\phi \to \infty$ and because, by Proposition 3-(i) in Alvarez and Lippi (2014), $\bar{y}$ is increasing in ψ and has range and domain $[0, \infty)$. For $\lambda = 0$ and $N(\Delta p_i) > 0$ we obtain: $\frac{\psi}{n} = B \; \frac{V(\Delta p)}{N(\Delta p_i)} \; \frac{n}{2\,(n+2)}$ .This follows from using the square root approximation of $\bar{y}$ for small $\psi \, (\lambda + r)^2$, the expression for $N(\Delta p_i) = n\sigma^2/\bar{y}$ and Proposition 1, i.e. $N(\Delta p_i) \, Var(\Delta p_i) = \sigma^2$. To obtain the expression for $\Psi(n, 0)$ we use Proposition 6 in Alvarez and Lippi (2014) where it is shown that for $\lambda = 0$ then $Kur\,(\Delta p_i) = 3n/(n+2)$.

D The CPI response to a monetary shock

To compute the IRF of the aggregate price level we analyze the contribution to the aggregate price level by each firm. Firms start with price gaps distributed according to g, the invariant distribution. Then the monetary shock displaces them, by subtracting the monetary shock δ to each of them. After that we divide the firms in two groups. Those that adjust immediately and those that adjust at some future time. Note that, for each firm in the cross section, it suffices to keep track only of the contribution to the aggregate price level of the first adjustment after the shock because the future contributions are all equal to zero in expected value.

Let $g\,(p; n, \lambda/\sigma^2, \bar{y})$ be the density of firms with price gap vector $p = (p_1, ..., p_n)$ at time $t = 0$, just before the monetary shock, which corresponds to the invariant distribution with constant money supply. The density g equals the density f of the steady state square norms of the price gaps given by Lemma 3 evaluated at $y = p_1^2 + \cdots + p_n^2$ times a correction factor:[31]

$$g\left(p_1, ..., p_n\,; n, \frac{\lambda}{\sigma^2}, \bar{y}\right) = f\left(p_1^2 + \cdots + p_n^2\,; n, \frac{\lambda}{\sigma^2}, \bar{y}\right) \frac{\Gamma\,(n/2)}{\pi^{n/2} \; (p_1^2 + \cdots + p_n^2)^{(n-2)/2}} \tag{37}$$

To define the impulse response we introduce two extra pieces of notation. First we let $\{(\bar{p}_1(t,p), \ldots, \bar{p}_n(t,p))\}$ the process for n independent BM, each one with variance per unit of time equal to σ^2, which at time $t = 0$ start at p, so $\bar{p}_i(0,p) = p_i$. We also define the stopping time $\tau(p)$, also indexed by the initial value of the price gaps p as the minimum of

[31] See Section 5 of Alvarez and Lippi (2014) for this result and the technical Appendix P for a derivation.

two stopping times, τ_1 and $\tau_2(p)$. The stopping time τ_1 denotes the first time since $t = 0$ that jump occurs for a Poisson process with arrival rate λ per unit of time. The stopping time $\tau_2(p)$ denotes the first time that $||\bar{p}(t,p)||^2 > \bar{y}$. Thus $\tau(p)$ is the first time a price change occurs for a firm that starts with price gap p at time zero. The stopped process $\bar{p}(\tau(0),p)$ is the vector of price gaps at the time of price change for such a firm.

The impulse response for the aggregate price level can be written as:

$$\mathcal{P}(t,\delta;\sigma,\lambda,\bar{y}) \;=\; \Theta(\delta;\sigma,\lambda,\bar{y}) \,+\, \int_0^t \theta(\delta,s;\sigma,\lambda,\bar{y})\,ds\,, \tag{38}$$

where $\Theta(\delta)$ gives the impact effect, the contribution of the monetary shock δ to the aggregate price level on impact, i.e. at the time of the monetary shock. The integral of the θ's gives the remaining effect of the monetary shock in the aggregate price level up to time t, i.e. $\theta(\delta,s)ds$ is the contribution to the increase in the average price level in the interval of times $(s, s+ds)$ from a monetary shock of size δ. Figure 3 displays several examples of impulse responses (the figures plots output, i.e. $(\delta - \mathcal{P})/\epsilon$). The functions θ and Θ are readily defined in terms of the density g, the process $\{\bar{p}\}$ and the stopping times τ:

$$\Theta(\delta;\sigma,\lambda,\bar{y}) \equiv \int_{||p(0)-\iota\delta||\geq\bar{y}} \left(\delta - \frac{\sum_{j=0}^n p_j(0)}{n}\right) g\left(p(0); n, \frac{\lambda}{\sigma^2}, \bar{y}\right)\, dp_1(0)\,\cdots\, dp_n(0)$$

and $\theta(\delta,t;\sigma,\lambda,\bar{y})$ is the density, i.e. the derivative with respect to t of the following expression:

$$\int_{||p(0)-\iota\delta||<\bar{y}} \mathbb{E}\left[-\frac{\sum_{j=0}^n \bar{p}_j\left(\tau(p),p\right)}{n}\,\mathbf{1}_{\{\tau(p)\leq t\}}\,\Big|\, p = p(0) - \iota\delta\right] g\left(p(0); n, \frac{\lambda}{\sigma^2}, \bar{y}\right) dp_1(0)\cdots dp_n(0)$$

where ι is a vector of n ones. This expression takes each firm that has not adjusted prices on impact, i.e. those with $p(0)$ satisfying $||p(0) - \iota\delta|| < \bar{y}$, weights them by the relevant density g, displaces the initial price gaps by the monetary shock, i.e. sets $p = p(0) - \iota\delta$, and then looks a the (negative) of the average price gap at the time of the first price adjustment, $\tau(p)$, provided that the price adjustment has happened before or at time t. We make 3 remarks about this expression. First, price changes equal the negative of the price gaps because price gaps are defined as prices minus the ideal price. Second, we define θ as a density because, strictly speaking, there is no effect on the price level due to price changes at *exactly* time t, since in continuous time there is a zero mass of firms adjusting at any given time. Third, we can disregard the effect of any subsequent adjustment because each of them has an expected zero contribution to the average price level. Fourth, the impulse response is based on the steady-state decision rules, i.e. adjusting only when $y \geq \bar{y}$ even after an aggregate shock occurs.

Given the results in Proposition 3 -Proposition 4 we can parametrize our model either in terms of $(n, \lambda, \sigma^2, \psi/B)$ or instead parametrize it, for each n, in terms of the implied observable statistics $(N(\Delta p_i), Std(\Delta p_i), \ell)$. These propositions show that this mapping is indeed one-to-one and onto. We refer to ℓ as an "observable" statistic, because we have shown that the "shape" of the distribution of price changes depends only on it.

Proposition 12 *Consider an economy whose firms produce n products and with steady state statistics $(N(\Delta p_i), Std(\Delta p_i), \ell)$. The cumulative proportional response of the aggregate price level $t \geq 0$ periods after a once and for all proportional monetary shock of size δ can be obtained from the one of an economy with one price change per period and with unitary standard deviation of price changes as follows:*

$$\mathcal{P}\left(t\,,\,\delta\,;\,N(\Delta p_i)\,,\,Std(\Delta p_i)\right) = Std(\Delta p_i)\;\mathcal{P}\left(t\,N(\Delta p_i)\,,\,\frac{\delta}{Std(\Delta p_i)}\,;\,1\,,\,1\right)\;. \tag{39}$$

This proposition extends the result of Proposition 8 in Alvarez and Lippi (2014) to the case of $\ell \equiv \lambda/N(\Delta p_i) > 0$.[32] The proof proceeds by verification. It is made of three parts. First we introduce a discrete-time, discrete-state version of the model. Second we show the scaling of time with respect to N_a , and finally the homogeneity of degree one with respect to $Std(\Delta p_i)$ and δ. The step by step passages are reported in the technical Appendix P.

The proposition establishes that the shape of the impulse response is completely determined by 2 parameters: n and ℓ, whose comparative static is explored in Figure 3. Economies sharing these parameters but differing in terms of $N(\Delta p_i)$ or $Std(\Delta p_i)$ are immediately analyzed by rescaling the values of the horizontal and/or vertical axis. In particular, a higher frequency of price adjustments will imply that the economy "travels faster" along the impulse response function (this is the sense of the rescaling the horizontal axis). Instead, the effect of a larger dispersion of price changes is seen by rescaling the monetary shock δ by $Std(\Delta p_i)$ and by a proportional scaling of the vertical axis. A further simplification to the last result is given by next corollary, showing that for small values of the monetary shocks one can overlook the scaling by $Std(\Delta p_i)$ so that, for a given n and ℓ determining the shape, the most important parameter is the frequency of price changes $N(\Delta p_i)$:

Corollary 1 *For small monetary shocks $\delta > 0$, the impulse response is independent of*

[32] The proof in Alvarez and Lippi is constructive in nature, exploiting results from applied math on the characterization of hitting times for brownian motions in hyper-spheres, which is not longer valid for $\lambda > 0$. Here we use a different strategy which relies on limits of discrete-time, discrete state approximations.

$Std(\Delta p_i)$. Differentiating equation (39) gives:

$$\mathcal{P}\left(t\,,\,\delta\,;\,N(\Delta p_i)\,,\,Std(\Delta p_i)\,\right) = \delta\,\frac{\partial}{\partial\delta}\,\mathcal{P}\left(t\,N(\Delta p_i)\,,\,0\,;\,1\,,\,1\,\right) + o(\delta)$$

for all $t > 0$ and, since $f(\bar{y}) = 0$, then the initial jump in prices can be neglected, i.e.:

$$\mathcal{P}\left(0\,,\,\delta\,;\,N(\Delta p_i)\,,\,Std(\Delta p_i)\,\right) \equiv \Theta_{n,\ell}\left(\delta; Std\left(\Delta p_i\right)\right) = o(\delta)\ .$$

E An economy with heterogenous sectors

Assume that there are S sectors, each with an expenditure weight $e(s) > 0$, and with different parameters so that each has $N(s)$ price changes per unit of time, and a distribution of price changes with kurtosis $Kur(s)$. In this case, after repeating the arguments above for each sector and aggregating, we obtain that the area under the IRF of *aggregate* output for a small monetary shock δ is

$$\mathcal{M}(\delta) \cong \delta\mathcal{M}'(0) = \frac{\delta}{6\,\epsilon}\sum_{s\in S}\frac{e(s)}{N(s)}Kur(s) = \frac{\delta}{6\,\epsilon}\,D\sum_{s\in S}d(s)\,Kur(s) \tag{40}$$

where D is the expenditure-weighted average duration of prices $D \equiv \sum_{s\in S}\frac{e(s)}{N(s)}$ and the $d(s) \equiv \frac{e(s)}{N(s)D}$ are weights taking into account both relative expenditures and durations. In the case in which all sectors have the same durations then $d(s) = e(s)$ and $\mathcal{M}$ is proportional to the kurtosis of the standardized data. Likewise, the same result applies if all sectors have the same kurtosis.[33] In general, if sectors are heterogenous in the durations (or expenditures), then the kurtosis of the sectors with longer duration (or expenditures) receive a higher weight in the computation of $\mathcal{M}$. For the French data, computation of the duration weighted kurtosis in equation (40) increases the estimated cumulative effect by about 15%, reflecting a correlation between the kurtosis and the duration of price changes.

F Frequency of price changes in Retail vs. Wholesale

In this appendix we document that wholesale prices are as sticky as retail prices for a broad cross section of products sold in grocery stores. For wholesale price we use PromoData, a dataset on manufacturer prices for packaged foods from grocery wholesalers (the largest

[33]The effect of heterogeneity in $N(\Delta p_i)$ on aggregation is well known, so that D is different from the average of $N(\Delta p_i)$'s, see for example Carvalho (2006) and Nakamura and Steinsson (2010).

Table 3: Weekly Frequency of Price Adjustment - Wholesale vs Retail Level

Data	Period	Frequency excl. Sales	Frequency
All Products			
PromoData (Wholesale)	2006-2012	0.09	0.14
IRI Symphony (Retail)	2001-2011	0.11	0.22
All Products (2006 - 2011)			
PromoData (Wholesale)	2006-2011	0.08	0.14
IRI Symphony (Retail)	2006-2011	0.12	0.23
Coffee			
PromoData (Wholesale)	2006-2012	0.17	0.20
IRI Symphony (Retail)	2001-2011	0.10	0.19
RMS	2006-2012	-	0.16

The table reports the weekly frequency of price adjustment using three datasets: Nielsen's PromoData, IRI Symphony, and Nielsen's Retail Scanner (RMS) data. The frequency of adjustment is computed at the product level and then aggregated across products using equal weights.

wholesaler in each location). PromoData provides the price per case charged by the manufacturer to the wholesaler for a UPC in a particular day, for 48 markets, over the period 2006-2012. The data includes information on almost 900 product categories and more than 500,000 UPC×Market products, and contain information on both base prices and "trade deals" (discounts offered to the grocery wholesalers to encourage promotions). We compute the frequency of price changes using base prices (excluding trade deals) as well as including trade deals.[34]

The frequency of price adjustment at the retail level is computed using the IRI Symphony data. The dataset contains weekly scanner price and quantity data covering a panel of stores in 50 metropolitan areas from January 2001 to December 2011, with multiple chains of retailers for each market. The dataset contains around 2.4 billion transactions from over 170,000 UPCs and around 3,000 stores. Goods are classified into 31 general product categories and a sales flag is provided when an item is on discount (thus we compute the frequency both including and excluding sales as in Section 2.2). To correct for measurement error (due to composition and time aggregation) we only retain price changes within the interval $0.1 \leq |\Delta p_i| \leq 100 \cdot \log(10/3)$. Finally, to compare with and extend Nakamura and Steinsson (2008), we compute the frequency of price changes for coffee using data on retail prices and

[34]In PromoData firms report only the dates in which their prices change. We thus assume that the price is constant between reporting dates. We discard the last price (uncompleted spell) and consider products with at least two price changes. The frequency of adjustment is computed at the weekly level for comparability with the retail data sets (even though our data may have a higher frequency). The frequency of adjustment is computed for each product (i.e. UPC x Market given that the data is not at the store level) and then aggregated using equal product weights.

sales from the Retail Scanner Data (RMS) by Nielsen. Our data is at the week-product-store level for the period of 2006-2013. The structure of the dataset is the same as the IRI Symphony data except that the RMS does not provide a sales flag, and covers about 200 cities.

Table 3 summarizes the main findings of this measurement exercise. The weekly frequency of price adjustment (sales excluded) for the entire wholesale data (PromoData) is 0.09 per week which compares with a mean frequency of adjustment of about 0.11 per week in the retail (IRI) data. Frequencies of comparable magnitude are detected across samples from different segments of the distribution chain, as well as for different items (coffee and beer, not reported) in the samples that exclude sales. Including sales makes the frequency of adjustment in retail somewhat higher than the frequency in wholesale.

G Simple special cases of Proposition 6

This section discusses some limiting cases in which tractable closed form expressions for the cumulative effect $\mathcal{M}$ as well as the frequency and kurtosis of price adjustments can be derived. The first two cases we illustrate assume either $n = 1$ or $n \to \infty$: we derive the implications for the cumulative output effect while considering the full range of values for $\ell \in (0, 1)$ and keeping the frequency of price changes constant. The last case restricts attention to $\ell = 0$ or $\ell = 1$ but allows for any value of $n \geq 1$.

G.1 Analytical computation of $\mathcal{M}$ in the case of $n = 1$

We give an analytical summary expression for the effect of monetary shocks in two interesting cases, those for one product, i.e. $n = 1$, and those for the large number of product, i.e. $n = \infty$. The summary expression is the area under the impulse response for output, i.e. the sum of the output above steady state after a monetary shock of size $\delta > 0$, which we denote as:

$$\mathcal{M}_n(\delta) = (1/\epsilon) \int_0^{\infty} [\delta - \mathcal{P}_n(\delta, t)] \, dt \tag{41}$$

where $1/\epsilon$ is related to the uncompensated labor supply elasticity and $\mathcal{P}_n(\delta, t)$ is the cumulative effect of monetary shock δ in the (log) of the price level after t periods. For large enough shocks, given the fixed cost of changing prices, the model display more price flexibility. Because of their prominence in the literature, and because of realism, we consider the case of small shocks δ by taking the first order approximation to equation (41), so we consider $\mathcal{M}_n(\delta) \approx \mathcal{M}'_n(0)\delta$.

For the case of $n = 1$ we obtain an analytical expression which, after normalizing by $N(\Delta p_i)$ depends only on $\lambda/N(\Delta p_i)$. Thus as $\lambda/N(\Delta p_i)$ ranges from 0 to 1 the model ranges from a version of the menu cost model of *Golosov and Lucas* to a version using *Calvo* pricing. The analytical expression is based upon the following characterization:

$$\mathcal{M}_1(\delta) = (1/\epsilon) \int_{-\bar{p}}^{\bar{p}-\delta} m(p_0)\, g(p_0 + \delta)\, dp_0 \tag{42}$$

where p_0 is the price gap *after* the monetary shocks and where $m(p)$ gives the contribution to the area under the IRF of firms that start with price gap, after the shock, equal to p_0. Since the monetary shock happens when the economy is in steady state, the distribution right after the shock has the steady state density h displaced by δ. Immediately after the shock the firms with the highest price gap have price gap $\bar{p} - \delta$. Note that the integral in equation (42) does not include the firms that adjust on impact, those that before the shock have price gaps in the interval $[-\bar{p}, \bar{p} - \delta)$, whose adjustment does not contribute to the IRF. The definition of m is:

$$m(p) = -\mathbb{E}\left[\int_0^\tau p(t)\, dt \;\middle|\; p(0) = p\right]$$

where τ is the stopping time denoting the first time that the firm adjusts its price. This function gives the integral of the negative of the price gap until the first price adjustment. This expression is based on the fact that those firms with negative price gaps, i.e. low markups, contribute positively to output being in excess of its steady state value, and those with high markups contribute negatively. Given a decision rule summarized by $\bar{p}$ we can characterize m as the solution to the following ODE and boundary conditions:

$$\lambda m(p) = -p + \frac{\sigma^2}{2} m''(p) \;\text{ for all } p \in [-\bar{p}, \bar{p}] \;\text{ and }\; m(p) = 0 \;\text{ otherwise .}$$

The solution for the function m is:

$$m(p) = -\frac{p}{\lambda} + \frac{\bar{p}}{\lambda}\left(\frac{e^{\sqrt{2\phi}\frac{p}{\bar{p}}} - e^{-\sqrt{2\phi}\frac{p}{\bar{p}}}}{e^{\sqrt{2\phi}} - e^{-\sqrt{2\phi}}}\right) \;\text{ for all }\; p \in [-\bar{p}\,,\, \bar{p}] \;.$$

$\phi \equiv \lambda \bar{p}^2/\sigma^2$. We then have:

$$\mathcal{M}(\delta) \approx \mathcal{M}'(0)\delta = (\delta/\epsilon) \int_{\bar{p}}^{\bar{p}} m(p)\, g'(p)\, dp = (\delta/\epsilon)\, 2 \int_0^{\bar{p}} m(p)\, g'(p)\, dp$$

since $m(\bar{p})g(\bar{p}) = 0$. The last equality uses that m is negative symmetric, i.e. $m(p) =$

$-m(-p)$, and that g is symmetric around zero. Using the expression for g in Section 3.1

$$g'(p) = -\frac{2\phi}{2\bar{p}^2\left(e^{\sqrt{2\phi}}-1\right)^2}\left(e^{\sqrt{2\phi}\left(2-\frac{p}{\bar{p}}\right)}+e^{\sqrt{2\phi}\frac{p}{\bar{p}}}\right) \quad \text{for} \quad p\in[0,\bar{p}] \ .$$

we obtain:

$$\delta\mathcal{M}'(0) \;=\; \left(\frac{\delta}{\epsilon}\right)\frac{-2}{\lambda\left(e^{\sqrt{2\phi}}-1\right)^2}\left(e^{\sqrt{2\phi}}\left(1+\phi-\frac{e^{\sqrt{2\phi}}+e^{-\sqrt{2\phi}}}{2}\right)\right) \quad .$$

Using the expression for $N(\Delta p_i)$ for the $n=1$ and simple algebra we can rewrite it as:

$$\delta\mathcal{M}'(0) \;=\; \left(\frac{\delta}{\epsilon}\right)\frac{1}{N(\Delta p_i)}\frac{e^{\sqrt{2\phi}}+e^{-\sqrt{2\phi}}}{\left(e^{\sqrt{2\phi}}+e^{-\sqrt{2\phi}}-2\right)^2}\left(e^{\sqrt{2\phi}}+e^{-\sqrt{2\phi}}-2-2\phi\right)$$

which yields the cumulative output effect of a small monetary shock of size δ.[35]

Kurtosis. We now verify that the expression can be equivalently obtained by computing the kurtosis, as stated in Proposition 6. For notation convenience let $x\equiv\sqrt{2\phi}$. Using the distribution of price changes derived in Section 3.1 and the definition of kurtosis we get

$$Kur\left(\Delta p_i\right) \;=\; \frac{2\ell\left(\frac{12}{x^4}-\frac{12+x^2}{x^2\left(e^{x/2}-e^{-x/2}\right)^2}\right)+1-\ell}{\left(2\ell\left(\frac{1}{x^2}+\frac{1}{2-e^{-x}-e^{x}}\right)+1-\ell\right)^2} \;=\; \frac{12-\frac{12x^2+x^4}{\left(e^{x/2}-e^{-x/2}\right)^2}+x^4\frac{1-\ell}{2\ell}}{2\ell\left(1+\frac{x^2}{2-e^{-x}-e^{x}}+x^2\frac{1-\ell}{2\ell}\right)^2}$$

Recall from Section 3.1 that $\ell=\frac{e^x+e^{-x}-2}{e^x+e^{-x}}$ so that , after some algebra

$$Kur\left(\Delta p_i\right) \;=\; 6\;\frac{e^x+e^{-x}}{\left(e^x+e^{-x}-2\right)^2}\quad\left(e^x+e^{-x}-2-x^2\right)$$

It is immediate that the kurtosis and the cumulative output effect satisfy Proposition 6.

[35] As a check of this formula compute the case for $\phi=0$, i.e. the cumulative output for the Golosov-Lucas model. In this case we let $\lambda=0$ and $\bar{p}>0$. In this case we have: $m(p)=-\frac{\bar{p}^2 p}{3\sigma^2}+\frac{p^3}{3\sigma^2}$. Also $g'(p)=-1/\bar{p}^2$ for $p\in(0,\bar{p}]$, so we have:

$$\mathcal{M}'(0)\delta=\left(\frac{\delta}{\epsilon}\right)\frac{2}{-3\sigma^2\bar{p}^2}\int_0^{\bar{p}}\left[-\bar{p}^2p+p^3\right]dp=\left(\frac{\delta}{\epsilon}\right)\frac{-2}{3\sigma^2\bar{p}^2}\left[-\frac{\bar{p}^4}{2}+\frac{\bar{p}^4}{4}\right]=\left(\frac{\delta}{\epsilon}\right)\frac{2\bar{p}^2}{3\sigma^2}\frac{2}{8}=\left(\frac{\delta}{\epsilon}\right)\frac{1}{N(\Delta p_i)}\frac{1}{6}$$

which is the same value obtained by taking the limit for $\phi\to 0$ in the general expression above.

G.2 Analytical computation of $\mathcal{M}$ in the case of $n = \infty$

Define

$$Y_n(t,\delta) \equiv \frac{1}{n}\sum_{i=1}^{n}[p_i(t)-\delta] = Y_n(t,0) - 2\delta\frac{\sum_{i=1}^{n} p_i(t)}{n} + \delta^2 \ .$$

where the $p_i(t)$ are independent of each other, start at $p_i(0) = 0$ and have normal distribution with $\mathbb{E}[p_i(t)] = 0$ and $Var[p_i(t)] = \sigma^2 t$. Then, by an application of the law of large numbers, we have:

$$Y_\infty(t,\delta) = Y_\infty(t,0) + \delta^2 = t\sigma^2 + \delta^2$$

Letting $\bar{Y} \equiv \lim_{n\to\infty} \bar{y}(n)/n$ we can represent the steady state optimal decision rule as adjusting prices when t, the time elapsed since last adjustment, attains $T = \bar{Y}/\sigma^2$. We compute the density of the distribution of products indexed by the time elapsed since the last adjustment t and, abusing notation, we denote it by f. This distribution is a truncated exponential with decay rate λ and with truncation T, thus the density is:

$$f(t) = \lambda\frac{e^{-\lambda t}}{1-e^{-\lambda T}} \ \text{ for all } \ t \in [0,T] \ .$$

The (expected) number of price changes per unit of time is given by the sum of the free adjustments and the ones that reach T, so

$$N(\Delta p_i) = \lambda + f(T) = \lambda\left[1 + \frac{e^{-\lambda T}}{1-e^{-\lambda T}}\right] = \frac{\lambda}{1-e^{-\lambda T}}$$

Note that, using the definition of T given above, $\lambda T = \bar{Y}\lambda/\sigma^2$ the parameter which indexes the shape of f and of the distribution of price changes. Since this figures prominently in this expressions we define:

$$\phi \equiv \lambda T = \frac{\bar{Y}\lambda}{\sigma^2} \ .$$

which is consistent with the definition of ϕ in Proposition 3. Using this definition we get:

$$\ell = \frac{\lambda}{N(\Delta p_i)} = 1 - e^{-\phi} \ \text{ and thus } \ N(\Delta p_i) = \frac{\lambda}{1-e^{-\phi}}$$

Impulse Response of Prices to a monetary Shock. We can now define the impulse response. Note that after the monetary shock firms that have adjusted their prices t periods ago, in average will adjust their price up by δ. This highlights that as $n \to \infty$ there is no selection.

Now we turn to the characterization of the impact effect Θ. In this case we have

$$Y_\infty(t,\delta) = Y_\infty(t,0) + \delta^2 = t\sigma^2 + \delta^2 \geq \bar{Y} = \sigma^2 T \iff t \geq T - \delta^2/\sigma^2 \ .$$

Thus the impact effect is:

$$\Theta(\delta) = \delta \int_{T-\delta^2/\sigma^2}^{T} f(t)dt = \delta\, \frac{e^{-\lambda T + \frac{\lambda}{\sigma^2}\delta^2} - e^{-\lambda T}}{1-e^{-\kappa}} = \delta\, \frac{e^{-\kappa + \frac{\lambda}{\sigma^2}\delta^2} - e^{-\kappa}}{1-e^{-\kappa}}$$

Using that $N(\Delta p_i) Var(\Delta p_i) = \sigma^2$ we can write:

$$\Theta(\delta) = \delta + \delta\, \frac{e^{-\kappa + \frac{\lambda}{N(\Delta p_i)}\frac{\delta^2}{Var(\Delta p_i)}} - 1}{1-e^{-\kappa}} = \delta + \delta\, \frac{\left(1 - \frac{\lambda}{N(\Delta p_i)}\right) e^{\frac{\lambda}{N(\Delta p_i)}\frac{\delta^2}{Var(\Delta p_i)}} - 1}{\lambda/N(\Delta p_i)}$$

Note that

$$\lim \Theta(\delta) = \begin{cases} \delta \left(\frac{\delta}{Std(\Delta p_i)}\right)^2 & \text{as } \lambda/N(\Delta p_i) \to 0 \\ 0 & \text{as } \lambda/N(\Delta p_i) \to 1 \end{cases}$$

and in general

$$\frac{\Theta(\delta)}{\partial(\lambda/N(\Delta p_i))} = \delta\, \frac{e^{\frac{\lambda}{N(\Delta p_i)}\frac{\delta^2}{Var(\Delta p_i)}} \left(\frac{\delta^2}{Var(\Delta p_i)}\frac{\lambda}{N(\Delta p_i)}\left(1 - \frac{\lambda}{N(\Delta p_i)}\right) - 1\right) + 1}{(\lambda/N(\Delta p_i))^2} < 0$$

whenever $\delta < 2\, Std(\Delta p_i)$.

$$\theta(t) \;=\; \delta e^{-\lambda t}\left[f\left(T - \delta^2/\sigma^2 - t\right) + \lambda \int_0^{T-\delta^2/\sigma^2 - t} f(s)ds \right] = \delta \frac{\lambda e^{-\lambda t}}{1-e^{-\lambda T}} \quad .$$

We can interpret $\theta(t)dt$ as $\theta(t)$ times the number of firms that adjust its price at times (t, dt). This is the sum of two terms. The first term is the fraction that adjust because they hit the boundary between t and $t + dt$. The second term is the fraction that have not yet adjusted times the fraction that adjust, λdt due to a free opportunity. Both terms are multiplied by $e^{-\lambda t}$ to take into account those firms that have received a free adjustment opportunity before after the monetary shock but before t.

Thus we have:

$$\mathcal{P}_\infty(t,\delta) \;=\; \Theta(\delta) + \delta \int_0^t \frac{\lambda e^{-\lambda s}}{1-e^{-\lambda T}} ds \;=\; \Theta(\delta) + \delta \frac{1 - e^{-\frac{\lambda}{N(\Delta p_i)}\, t\, N(\Delta p_i)}}{\lambda/N(\Delta p_i)}$$

Using $\mathcal{P}_\infty$ we can compute the IRF for output, and a summary measure for it, namely the area below it:

$$\mathcal{M}_\infty(\delta) \;=\; \frac{1}{\epsilon}\int_0^T \left[\delta - \mathcal{P}_\infty(\delta, t)\right] dt \;\approx\; \frac{\delta}{\epsilon\, N(\Delta p_i)} \left[\frac{1-(1+\phi)\, e^{-\phi}}{(1-e^{-\phi})^2}\right]$$

where the approximation uses the expression for small δ, i.e. its first order Taylor's expansion.

Kurtosis. For completeness we also include here an expression for the kurtosis of the distribution of price changes in the case of $n = \infty$. Price changes are distributed as:

$$\mathbb{E}\left[(\Delta p_i)^2\right] \;=\; \sigma^2/N(\Delta p_i) = \frac{\sigma^2}{\lambda}\frac{\lambda}{N(\Delta p_i)} = \frac{T\sigma^2}{T\lambda}\frac{\lambda}{N(\Delta p_i)} = T\sigma^2\,\frac{1}{T\lambda}\frac{\lambda}{N(\Delta p_i)}$$

$$\begin{aligned}
\mathbb{E}\left[(\Delta p_i)^4\right] \;&=\; 3\frac{\lambda}{N(\Delta p_i)}\int_0^T \frac{(\sigma^2 t)^2 \lambda e^{-\lambda t}}{1-e^{-\lambda T}} dt + \left(1 - \frac{\lambda}{N(\Delta p_i)}\right) 3\left(\sigma^2 T\right)^2 \\
&=\; 3\sigma^4 T^2 \left[\frac{2 - e^{-\lambda T}\left(\lambda T\left(\lambda T + 2\right) + 2\right)}{(T\lambda)^2} + \left(1 - \frac{\lambda}{N(\Delta p_i)}\right)\right]
\end{aligned}$$

Some algebra shows that kurtosis is then given by:

$$\frac{\mathbb{E}\left[(\Delta p_i)^4\right]}{\left(\mathbb{E}\left[(\Delta p_i)^2\right]\right)^2} \;=\; 6\,\frac{1 - e^{-\phi}\,(1+\phi)}{(1-e^{-\phi})^2}$$

It is immediate to use the expressions above to verify Proposition 6.

G.3 Analytical computation for $\ell = 0$ or $\ell = 1$ (any n).

For $\ell = 0$, or equivalently $\lambda = 0$, we use the result in Alvarez and Lippi (2014) for

$$\mathcal{T}_{n+2}(y) = \frac{\bar{y} - y}{(n+2)\sigma^2}$$

gives:

$$\mathcal{M}'(0) \;=\; \frac{1}{n\epsilon}\int_0^{\bar{y}} \left[\frac{n(\bar{y}-y) - 2y}{(n+2)\sigma^2}\right] f(y)\, dy$$

and using the following expression for f from Alvarez and Lippi (2014) :

$$f(y) = \frac{1}{\bar{y}} [\log(\bar{y}) - \log(y)] \text{ if } n = 2, \text{ and}$$
$$f(y) = (\bar{y})^{-\frac{n}{2}} \left(\frac{n}{n-2} \right) \left[(\bar{y})^{\frac{n}{2}-1} - (y)^{\frac{n}{2}-1} \right] \text{ otherwise}$$

gives that:

$$\mathcal{M}'(0) = \frac{1}{n\epsilon} \frac{2\,\bar{y}\,n(n-2)}{(n^2-4)\sigma^2} = \frac{1}{\epsilon} \frac{Kurt(\Delta p_i)}{6\,N(\Delta p_i)}$$

which verifies the equality in Proposition 6.

For $1 < n < \infty$ and $\ell = 1$, with $\lambda > 0$ and $\sigma^2 > 0$, using Proposition 3 it must be the case that $\bar{y} = \infty$. In this case, $N(\Delta p_i) = \lambda \ell = \lambda$, and the distribution of price changes is independent across each of the n products, and given by a Laplace distribution, which has kurtosis 6. Likewise $\mathcal{T}_{n+2}(y) = 1/\lambda$ for all $y \geq 0$. Thus, using equation (26) we obtain the desired result.

TECHNICAL APPENDICES

The real effects of monetary shocks in sticky price models: a sufficient statistic approach

F. Alvarez , H. Le Bihan, and F. Lippi

April 27, 2016

H Data Appendix

This appendix provides further empirical evidence. Section H.1 offers more summary statistics about the French CPI and a robustness-to-trimming exercise. Section H.2 explores the extent to which the statistical protocols used to measure prices are responsible for the small price changes, as suggested by Eichenbaum et al. (2014).

H.1 More statistics on the French CPI and robustness to trimming

Table 4 reports the frequency of price changes as well as selected moments of the distribution of price changes. The basic patterns that emerge form the CPI data (frequency of price change, average and standard deviation of price changes) match those documented by Berardi, Gautier, and Le Bihan (2015) for France and are representative of those obtaiend by Alvarez et al. (2006) for the Euro area. With the qualification that the frequency of price changes is typically found to be smaller in the Euro area than in the US, they also broadly match the US evidence by e.g. Nakamura and Steinsson (2008). The frequency of price change is around 17% per month, or about 2 price changes per year. The fraction of price decreases among price changes is around 40%. The average *absolute* price change (not reported in the table) is sizable (9.2%), as is the standard deviation of price changes (16.6%).

A second investigation on measurement error was developed by varying the upper and lower thresholds of small and large price changes used to define outliers. Results are displayed in Table 5. In each of the variants considered in Table 5, both kurtosis and the fraction of small price changes remain large. The lowest level of kurtosis obtains when we use the most stringent thresholds for outliers. If in the vein Eichenbaum et al. (2014) we consider that all price changes lower in absolute value that 1 percent are presumably measurement errors, and discard them from the sample, the resulting kurtosis is reduced to 7.1. Furthermore, some large price changes may as well be measurement (transcription) errors. If we restrict the sample to price changes larger in absolute value that 1 percent, and lower than log(2) (meaning that observaions with multiplied by more than 2, or divided by more than 2,– a stricter threshold than the 99^{th} percentile– are considered as likely measurement errors), the kurtosis is further reduced to 6.2.

H.2 Small price changes and measurement errors

This appendix examines to what extent the arguments of Eichenbaum et al. (2014) apply to our data and investigates the robustness of our findings to various criteria for trimming the data. Measurement errors may arise for several reasons. Eichenbaum, Jaimovich, and Rebelo

Table 4: Selected moments from the size distribution of price changes

	Dominick's DFF	CPI Data France all records	CPI Data France exc.sales
Frequency of price changes	32.7	17.1	15.7
Moments for the size of price changes: Δp_i			
Average	0.1	0.3	3.2
Standard deviation	24.5	15.6	11.9
Moments of standardized price changes: z			
Kurtosis	4.0	8.0	8.9
Number of obs. with $\Delta p_i \neq 0$	295,692	1,530,878	1,266,507

Source for Dominick's is Kilts Center for Marketing, data are weekly scanner price records for 400 weeks from 1989 to 1997. Source for CPI is INSEE, monthly price records from French CPI, data from 2003:4 to 2011:4. Coverage is around 65% of CPI weight since rents, and prices of fresh food and centrally collected items (e.g. electricity, train and airplane tickets) are not included in the dataset. Frequency of price change is the average fraction of price changes per week (Dominick's) or month (INSEE) , in percent. Size of price change is the first-difference in the logarithm of price per unit, expressed in percent. Observations with imputed prices or quality change are discarded. Observations outside the interval $0.1 \leq |\Delta p_i| \leq P99$ are removed as outliers. "Exc. sales" exclude observations flagged as sales by the INSEE data collectors. Moments are computed aggregating all prices changes using CPI weights at the product level. The "kurtosis" row report kurtosis for the standardized price change $z_{ijt} = \frac{\Delta p_{ijt} - m_j}{\sigma_j}$ where m_j and σ_j are the mean and standard deviation of price changes in category j (see the text).

(2011) and Eichenbaum et al. (2014) articulate two concerns about the small price change. First they notice that in scanner data studies the price level of an item is typically computed as the ratio of recorded weekly revenues to quantity sold. To the extent that there are temporary or individual specific discounts (say coupons), this will generate spurious small price changes.[36] Moreover Eichenbaum et al. (2014) highlight a related problem for some CPI items: they spot 27 items (named ELIS in the BLS terminology) that are problematic because these prices are typically computed as a Unit Value Index (a ratio of expenditure to quantity purchased), or they are not consistently recorded in the same outlet, or they are the price of a bundle of goods (for instance the sum of airplane fare and airport tax). We were able to match these items with their counterparts in our French dataset. Out of the 27 problematic items 15 are not present in our data because in the French CPI those items are not recorded by a field agent but are centrally collected (thus not made available

[36] Notice that in principle CPI data are immune from this type of measurement error, as these data are direct transaction prices observed by a field agent. Indeed, in the instance of a temporary discount, the CPI dataset will record either no price change, or the large price change of observed during the discount, if the field agent happens to be collecting data during the temporary discount. Further, the protocol of data collection requires that the field agent records the price faced by a regular customer, not benefiting from individual-specific discounts.

Table 5: Robustness to trimming

Type of trimming	Freq	Kur(z)	Std(dp)	N obs	case
$\|\Delta p\| < P99$	17.05	7.98	15.63	1530878	0
exc sales	15.71	8.85	11.88	1266507	1
<log(10/3)	17.09	8.89	16.60	1542527	2
<log(2)	17.08	8.53	16.12	1537895	3
<100pct	16.86	7.03	13.33	1482044	4
>0	17.17	8.12	15.56	1540243	5
>0.5pct	16.63	7.61	15.86	1480753	6
>1pct	15.34	7.07	16.65	1377209	7
>1pct <log(2)	15.14	6.15	14.21	1328378	8
>1pct nosales	13.95	7.84	12.68	1116076	9

Source is INSEE, monthly price records from French CPI, data from 2003:4 to 2011:4. Coverage is aroud 65% of CPI weight since rents, and prices of fresh food and centrally collected items (e.g. electricity, train and airplane tickets) are not included in the dataset. Freq. denotes monthly frequency of price change in percent. Size of price change Δp are the first-difference in the logarithm of price per unit, expressed in percent.Std(dp) is standard devation of log price change. Kur$[z]$ denotes kurtosis of the distribution of standardized price changes. Standardized price changes are computed at the category of good * type of outlet level. Observations with imputed prices or quality change are discarded. Moments are computed aggregating all prices changes using CPI weigths at the product level. Each row describes a sub-sample constructed applying the filter described by the column "type of trimming". The subsample with flag "case 0" is the baseline sample in the main text of the paper: price changes are included if they are larger in absolute value than 0.1 percent, and lower in absolute value than the 99^{th} percentile; sales are included. Each subsequent row describes the impact of changing one (or two) of these thresholds and criteria, the one(s) that is explicitely mentionned. For example the second row considers the sample with $|\Delta p| > 0.1$, $|\Delta p| < P99$,and sales excluded; the third row considers the sample with $|\Delta p| > 0.1$, $|\Delta p| < log(10/3)$,and sales included "Ex. sales" exclude observations flagged as sales by the INSEE data collectors.

in the subset of CPI we have access to).[37] Concerning the 12 remaining items virtually no price record in the French CPI is computed as a Unit Value Index, which is hypothesized by Eichenbaum et al. (2014) as a major source of small price changes. Inspecting the patterns of price changes over these 12 potentially "problematic" items in our dataset shows that the amount of small price changes is not significantly different from the one detected over the rest of our sample. One exception is the price of "Residential water" where it can be suspected that many small variations in local taxes occur.[38]

Finally, Table 6 compares the fraction of small price changes in US vs the French data.

[37] These items are Hospital room in-patient; Hospital in-patient services other than room ; Electricity; Utility natural gas service; Telephone services, local charges ; Interstate telephone services ; Community antenna or cable TV ; Cigarettes; Garbage and trash collection; Airline fares; New cars; New trucks; Ship fares; Prescription drugs and medical supplies; Automobile insurance.

[38] Otherwise, on the bulk of consumption items, there are no local taxes in France, and the main, nation-wide, rate of the Value Added Tax rate did not move over the sample period.

Table 6: Fraction of small price changes: US and French CPI

Moments for the absolute value of price changes: $\lvert\Delta p_i\rvert$				
	France	US	Normal	Laplace
Average $\lvert\Delta p_i\rvert$	9.2	14.0		
Fraction of $\lvert\Delta p_i\rvert$ below 1%	11.8	12.5		
Fraction of $\lvert\Delta p_i\rvert$ below 2.5%	32.5	24.0		
Fraction of $\lvert\Delta p_i\rvert$ below 5%	57.1	40.6		
Fraction of $\lvert\Delta p_i\rvert$ below $(1/14) \cdot \mathbb{E}(\lvert\Delta p_i\rvert)$	2.4	12.5	4.5	6.9
Fraction of $\lvert\Delta p_i\rvert$ below $(2.5/14) \cdot \mathbb{E}(\lvert\Delta p_i\rvert)$	13.5	24.0	11.3	16.4
Fraction of $\lvert\Delta p_i\rvert$ below $(5/14) \cdot \mathbb{E}(\lvert\Delta p_i\rvert)$	28.7	40.6	22.4	30.0
Number of obs	1,542,586	1,047,547		

Note: For France the source is INSEE monthly price records from the French CPI (2003:4 to 2011:4). Coverage is around 65% of CPI weight since rents, and prices of fresh food and centrally collected items (e.g. electricity, train and airplane tickets) are not included in the dataset. Frequency of price change is monthly, in percent. Size of price change are the first-difference in the logarithm of price per unit, expressed in percent. Data are trimmed as in the baseline of Table 4. Observations with imputed prices or quality change are discarded. Moments are computed aggregating all prices changes using CPI weights at the product level. The US data are taken from Eichenbaum et al. (2014) Table 1, and refer to "Posted price changes" from 1998:1 to 2011:6. The mean absolute size of price changes is taken from Klenow and Kryvtsov (2008) table III where data are from 1998:1 to 2005:1. Figures for the US are weighted and cover around 70% of the CPI (US CPI includes owners equivalent rents, while French CPI does not). In the third panel we compute the threshold for defining small price changes as fraction of the mean so as to match the US figures in column 2 of the second panel. The Normal and Laplace distributions used in the last two columns have a zero mean and, without loss of generality, standard deviation equal to one.

The table uses the same thresholds of Eichenbaum et al. (2014) to measure the fraction of small price changes. The presence of small price changes (in absolute value) is at first sight a more prominent fact in France than in the US. One factor that may contribute to explaining this pattern is the fact that sales are less prevalent in France. Measurement error, as discussed above, may play a role. We nevertheless observe that, if we define small price changes as relative to the mean average price change, rather than with an absolute threshold, the fraction of small price change appears to be lower in France than in the US, as shown in Table 6.

I Details of the solution for the model with $n = 1$

Integrating the Bellman equation gives the following value function

$$V(p) = \frac{Bp^2 + \lambda V(0)}{\lambda + r} + \frac{B\sigma^2}{(\lambda + r)^2} + C\left(e^{p\sqrt{\frac{2(\lambda+r)}{\sigma^2}}} + e^{-p\sqrt{\frac{2(\lambda+r)}{\sigma^2}}} \right)$$

where we already used that $V(p) = V(-p)$. Notice that the value function has a minimum (and zero derivative) at $p = 0$, which is the optimal return point. The constant C and the threshold value $\bar{p}$ are the values that solve the 2 equation system given by the value matching condition and the smooth pasting conditions.

The expected time to adjustment, $T(p)$ obeys the differential equation $\lambda T(p) = 1 + \frac{\sigma^2}{2} T''(p)$ with boundary condition $T(\bar{p}) = 0$. Given the symmetry of the law of motion for p, the function is symmetric, i.e. $T(p) = T(-p)$. Integrating gives $T(p) = \frac{1}{\lambda}\left(1 - \frac{e^{\sqrt{\frac{2\lambda}{\sigma^2}}p} + e^{-\sqrt{\frac{2\lambda}{\sigma^2}}p}}{e^{\sqrt{\frac{2\lambda}{\sigma^2}}\bar{p}} + e^{-\sqrt{\frac{2\lambda}{\sigma^2}}\bar{p}}}\right)$.

The distribution of price gaps $g(p)$ satisfies the Kolmogorov forward equation $0 = -\frac{2\lambda}{\sigma^2} g(p) + g''(p) \quad \text{for} \quad 0 < |p| \leq \bar{p}$. The density is symmetric, $g(p) = g(-p)$, and satisfies the boundary conditions: $g(\bar{p}) = 0$ and it integrates to one i.e. $2\int_0^{\bar{p}} g(p)\, dp = 1$ where we used that it is symmetric.[39]

J Proof that $\lim_{\bar{y}\to\infty} \xi(\sigma^2, r+\lambda, n, \bar{y}) = 0$

Note that, by examining the definition of κ_i and the sums in the expression for ξ we have that:

$$\lim_{\bar{y}\to\infty} \xi(\sigma^2, r+\lambda, n, \bar{y}) = \lim_{\bar{y}\to\infty} \xi\left(1, 1, n, \frac{(r+\lambda)\,\bar{y}}{\sigma^2}\right)$$

so this limit cannot depend on $r+\lambda$ or σ^2. Thus we denote it as:

$$\bar{\xi}(n) \equiv \lim_{\bar{y}\to\infty} \xi(1, 1, n, \bar{y})$$

So we have:

$$\bar{y} \approx \frac{\psi}{B}(r+\lambda)\left[1 - \bar{\xi}(n)\right] \quad \text{for large } \psi\,.$$

Now we show that $\bar{\xi}(n) = 0$. First we notice that the power series:

$$g(x) = \sum_{i=1}^{\infty} \prod_{s=1}^{i} \frac{1}{(s+2)(n+2s+2)}\, x^i$$

[39] The first boundary can be derived as the limit of the discrete time, discrete state, low of motion where each period is of length Δ and where p increases or decreases with probability $1/2$, so that $g(p) = \frac{1}{2}g(p+\Delta) + \frac{1}{2}g(p-\Delta)$. At the boundary $\bar{p}$ this law of motion is $g(\bar{p}) = \frac{1}{2}g(\bar{p}-\Delta)$, which shows that $g(\bar{p}) \downarrow 0$ as $\Delta \downarrow 0$.

converges for all values of x since its coefficients satisfy the Cauchy-Hadamard inequality. Then we can write:

$$\xi(1,1,n,\bar{y}) \equiv \frac{\frac{2(n+2)}{\bar{y}}\frac{1}{g(\bar{y})} + \frac{1}{g(\bar{y})} + \frac{1}{\bar{y}^2}}{\frac{2(n+2)}{\bar{y}}\frac{1}{g(\bar{y})} + 2\frac{1}{g(\bar{y})} + \sum_{i=1}^{\infty}\omega(i,\bar{y})\,(2+i)}$$

where the weights $\omega(i,\bar{y})$ are given by:

$$\omega(i,x) = \frac{\frac{x^i}{\prod_{s=1}^{i}(s+2)(n+2s+2)}}{\sum_{j=1}^{\infty}\prod_{s=1}^{j}\frac{1}{(s+2)(n+2s+2)}\;x^j}$$

Note that for higher x the weights of smaller i decrease relative to the ones for higher i. Now since $g(\bar{y}) \to \infty$ as $\bar{y} \to \infty$, then:

$$\bar{\xi}(n) = \frac{1}{\lim_{\bar{y}\to\infty}\sum_{i=1}^{\infty}\omega(i,\bar{y})\,(2+i)}$$

To show that $\bar{\xi}(n) = 0$, suppose, by contradiction that is finite. Say, without loss of generality that equals $j+2$ for some integer j. Note that, by the form of the $\omega's$ and because $g(\bar{y})$ diverges as $\bar{y}$ gets large enough, then by any j and $\epsilon > 0$ there exist a y^* large enough so that $\sum_{i=1}^{j}\omega(i,\bar{y}) < \epsilon$ for any $\bar{y} > y^*$. Thus, the expected value must be larger than $2+j$.

Finally, we consider the case of $n \to \infty$. In this case we have that, the value function divided by n gives:

$$v = \min_T B\int_0^T \sigma^2\, t\, e^{-(\lambda+r)}\, dt + e^{-(r+\lambda)T}(\Psi + v)$$

where $\Psi = \lim_{n\to\infty}\psi/n$. The first order condition for T gives, for a finite T:

$$0 = \big(B\,\sigma^2\, T - (r+\lambda)\Psi\big) - (r+\lambda)e^{-(r+\lambda)T}v \tag{43}$$

Now consider the case where $\Psi \to \infty$. Note that v is finite since $T = \infty$, a feasible strategy as a finite value. Also let $\bar{Y} = \sigma^2 T = \lim_{n\to\infty}\frac{\bar{y}(n)}{n}$. Note that as $\Psi \to \infty$ then $\bar{Y}$ must also diverge towards ∞. Dividing the previous expression by Ψ:

$$\frac{\bar{Y}}{\Psi} = \frac{(r+\lambda)}{B} + (r+\lambda)e^{-(r+\lambda)T}\frac{v}{\Psi}$$

and taking the limits:

$$\lim_{\Psi\to\infty}\frac{\bar{Y}}{\Psi} = \frac{r+\lambda}{B} \quad . \quad \square$$

K Note on solutions of value function $v(y)$, expected time to adjust $\mathcal{T}(y)$ and invariant density of the squared price gap $f(y)$.

First we state a proposition which gives an explicit closed form solution to the value function $v(y)$ in the inaction region, i.e. for $y \in (0, \bar{y})$ subject to $v(0) < \infty$. The solution is parameterized by $\beta_0 = v(0)$.

Proposition 13 *Let $\sigma > 0$. The ODE in equation (5) is solved by the analytical function: $v(y) = \sum_{i=0}^{\infty} \beta_i \, y^i$, for $y \in [0, \bar{y}]$ where, for any β_0, the coefficients $\{\beta_i\}$ solve: $\beta_0 = \frac{n\sigma^2}{r}\beta_1$, $\beta_2 = \frac{(r+\lambda)\beta_1 - B}{2\sigma^2(n+2)}$, $\beta_{i+1} = \frac{r+\lambda}{(i+1)\sigma^2 \ (n+2i)} \ \beta_i$ for $i \geq 2$.*

The function described in this proposition allows to fully characterize the solution of the firm's problem. One can use it to evaluate the two boundary conditions described above, value matching and smooth pasting, and define a system of two equations in two unknowns, namely β_0 and $\bar{y}$.

The alert reader may have noticed that to solve for the invariant density f we have followed a standard procedure, i.e. set a 2nd order ordinary linear difference equation (the Kolmogorov forward equation) and find its solutions in terms of two constant, and using two boundary conditions to find the value of the constants. Instead to solve for v and $\mathcal{T}$ we have followed a different approach, we guess an infinite expansion around $y = 0$ and compute its coefficients. Additionally, it may have looked that we did not provide enough boundary conditions to be able to solve for $\mathcal{T}$ and v. For instance, for $\mathcal{T}$ we gave only one equation as boundary conditions, namely $\mathcal{T}(\bar{y}) = 0$. Here we explain that we could have followed the more standard route, which required an analysis of the behavior close to the $y = 0$ boundary, to set one constant to zero and also would have produced a less informative result, i.e. one in terms of modified Bessel functions. Nevertheless we include it here for completeness.

Note that $v(y)$, $\mathcal{T}(y)$ and $f(y)$ are solutions to a linear ODE on y whose homogeneous component, say $q(\cdot)$, solves :

$$y \, q''(y) + a \, q'(y) + b \, q(y) \ = \ 0 \tag{44}$$

for $y \in [0, \bar{y}]$, for (different) constants a and b, with different particular solution, and different boundary conditions. The general solution of the homogeneous equation (44) is given by:

$$q(y) = |b \, y|^{(1-a)/2} \left[C_1 I_\nu \left(2\sqrt{|b \, y|} \right) + C_2 K_\nu \left(2\sqrt{|b \, y|} \right) \right] \tag{45}$$

provided that $b\,y < 0$, i..e. that $b < 0$, where C_1 and C_2 are arbitrary constants, $\nu = |1-a|$ and where I_ν and K_v are the modified Bessel functions of the first and second kind respectively. The values of $b = -\lambda/(2\sigma^2)$ in the three cases. The value of $a = n/2$ for $\mathcal{T}$ and for v, which are the same Kolmogorov backward equation, and $a = -(n/2 - 2)$ for f, which is the Kolmogorov forward equation.

It is important to notice the behavior of $I_\nu(z)$ and $K_\nu(z)$ for values of $0 < z$ but very close to zero. We have:

$$I_\nu \backsim \frac{1}{\Gamma(\nu+1)}\left(\frac{z}{2}\right)^\nu \tag{46}$$

and

$$K_\nu \backsim \begin{cases} \frac{\Gamma(\nu+1)}{2}\left(\frac{2}{z}\right)^\nu & \text{if } \nu > 0 \\ -\log(z/2) - \gamma & \text{if } \nu = 0 \end{cases} \tag{47}$$

We thus have that each of the solution will behave as:

$$\begin{aligned} I_{|1-a|}\left(y^{1/2}\right)\, y^{(1-a)/2} \quad &\backsim \quad \frac{1}{\Gamma(|1-a|+1)}\left(\frac{y^{1/2}}{2}\right)^{|1-a|} \quad y^{(1-a)/2} \\ &= \quad \frac{1}{\Gamma(|1-a|+1)}\left(\frac{1}{2}\right)^{|1-a|} \quad y^{(1-a)/2+|1-a|/2} \end{aligned}$$

So if $1-a = -|1-a|$, i.e. if $1-a \leq 0$, the value of this product is finite at $y \downarrow 0$. Otherwise it diverges to ∞. Likewise for $\nu = |1-a| > 0$:

$$\begin{aligned} K_{|1-a|}\left(y^{1/2}\right)\, y^{(1-a)/2} \quad &\backsim \quad \frac{\Gamma(|1-a|+1)}{2}\left(\frac{2}{y^{1/2}}\right)^{|1-a|} \quad y^{(1-a)/2} \\ &= \quad \frac{\Gamma(|1-a|+1)}{2}\left(\frac{2}{1}\right)^{|1-a|} \quad y^{(1-a)/2-|1-a|/2} \end{aligned}$$

So if $1-a = |1-a|$, i.e. if $1-a \geq 0$, the value of this product is finite at $y \downarrow 0$. Otherwise it diverges to ∞. The case of $\nu = 0$ i.e. $a = 1$ is special, but $K_0(z)$ also diverges and $I_0(z)$ converges to a non-zero constant as $z \downarrow 0$.

Note that $v(0)$ and $\mathcal{T}(0)$ are both finite. For these two cases the Kolmogorov backward equation has $a = n/2$ so $1-a \geq 0$ iff $n \geq 2$. In these cases we have that C_2, the constant associated with K_ν must be zero. We can use the constant C_1 to impose the boundary condition $\mathcal{T}(\bar{y}) = 0$ for $\mathcal{T}$ and to have a one dimensional representation of v in the range of inaction given $\bar{y}$. Then we can use smooth pasting and value matching, i.e. two boundary conditions, to find the constants C_1 and $\bar{y}$.

Note that for f we don't require that $f(0)$ be zero, since the density at zero gap can be infinite if the y mean reverts to zero fast enough. Thus in this case we will, in general, have both constants be non-zero.

L Power series representation of Kurtosis

Given $(\lambda, \sigma^2, \bar{y})$ the kurtosis of the steady state price distribution can be written as:

$$Kur\left(\Delta p_i\right) = \frac{Q(0)}{\frac{\sigma^4}{N(\Delta p_i)^2}} = \frac{\left(\lambda/\sigma^2\right)^2 Q(0)}{\left(\mathcal{L}(\phi, n)\right)^2}$$

where $Q(y)$ is the expected fourth moment at the time of adjustment τ conditional on having today a squared price gap y, i.e.

$$Q(y) = \mathbb{E}\left(\Delta p_i^4(\tau) \,|\, y(0) = y\right) = \frac{3}{(n+2)n}\, \mathbb{E}\left(y^2(\tau)|y(0) = y\right)$$

where $y(\tau)$ is the value of the squared price gap at the stopping time and where, using results from Alvarez and Lippi (2014), we have that $Kur(\Delta p_i|y) = \frac{3n}{(n+2)}$ and the variance is $Var\left(\Delta p_i \,|\, ||p||^2 = y\right) = y/n$. Notice that for $y \in [0, \bar{y}]$ the function $Q(y)$ obeys the o.d.e.:

$$\lambda Q(y) = \lambda \frac{3y^2}{(n+2)n} + Q'(y)\, n\sigma^2 + Q''(y)\, 2\sigma^2 y$$

with boundary condition $Q(\bar{y}) = \frac{3\bar{y}^2}{(n+2)n}$. Assuming that $Q(y) = \sum_{i=0}^{\infty} a_i y^i$, matching coefficients, and writing them as function of a_0 one obtains:

$$\begin{aligned} a_1(a_0) &= \frac{a_0}{\frac{\sigma^2}{\lambda} n}\ ,\ a_2(a_0) = \frac{a_1(a_0)}{2\frac{\sigma^2}{\lambda}(n+2)}\ ,\ a_3(a_0) = \frac{a_2(a_0) - \frac{3}{(n+2)n}}{3\frac{\sigma^2}{\lambda}(n+4)} \text{ and} \\ a_{i+1}(a_0) &= \frac{a_i(a_0)}{(i+1)\frac{\sigma^2}{\lambda}(n+2i)} \text{ for } i \geq 3 \end{aligned}$$

Thus $Q(0) = a_0$ is determined as the solution to $\sum_{i=0}^{\infty} a_i(a_0)\, \bar{y}^i = \frac{3\,\bar{y}^2}{(n+2)n}$. After tedious but simple algebra this gives:

$$Q(0) = a_0 = \frac{3n}{(n+2)} \left(\frac{\sigma^2}{\lambda}\right)^2 \left[\frac{\phi^2 + \sum_{i=3}^{\infty}\left(\prod_{j=3}^{i} \frac{n}{j[n+2(j-1)]}\right)\phi^i}{1 + \sum_{i=1}^{\infty}\left(\prod_{j=1}^{i} \frac{n}{j[n+2(j-1)]}\right)\phi^i}\right]$$

where $\phi \equiv \frac{\lambda \bar{y}}{n\sigma^2}$.

Replacing $Q(0)$ into $Kur\left(\Delta p_i\right) = \frac{\left(\lambda/\sigma^2\right)^2 Q(0)}{\mathcal{L}^2}$ and using equation (6) for $\mathcal{L}(\phi, n)$ we get

$$Kur\left(\Delta p_i\right) = \frac{3n}{(n+2)} \left[\frac{\phi^2 + \sum_{i=3}^{\infty}\left(\prod_{j=3}^{i} \frac{n}{j[n+2(j-1)]}\right)\phi^i}{1 + \sum_{i=1}^{\infty}\left(\prod_{j=1}^{i} \frac{n}{j[n+2(j-1)]}\right)\phi^i}\right] \left(\frac{1 + \sum_{i=1}^{\infty}\left(\prod_{j=1}^{i} \frac{n}{j[n+2(j-1)]}\right)\phi^i}{\sum_{i=1}^{\infty}\left(\prod_{j=1}^{i} \frac{n}{j[n+2(j-1)]}\right)\phi^i}\right)^2$$

Thus

$$Kur\left(\Delta p_i\right) = \frac{3n}{(n+2)} \frac{\left(\phi^2 + \sum_{i=3}^{\infty}\left(\prod_{j=3}^{i} \frac{n}{j[n+2(j-1)]}\right)\phi^i\right)\left(1 + \sum_{i=1}^{\infty}\left(\prod_{j=1}^{i} \frac{n}{j[n+2(j-1)]}\right)\phi^i\right)}{\left(\sum_{i=1}^{\infty}\left(\prod_{j=1}^{i} \frac{n}{j[n+2(j-1)]}\right)\phi^i\right)^2} \tag{48}$$

For future reference note that $Kur\left(\Delta p_i\right)/N\left(\Delta p_i\right) = (1/\lambda)\mathcal{L}(\phi, n)\, Kur\left(\Delta p_i\right)$ so

$$\frac{Kur\left(\Delta p_i\right)}{N\left(\Delta p_i\right)} = \frac{1}{\lambda}\frac{3n}{(n+2)} \frac{\left(\phi^2 + \sum_{i=3}^{\infty}\left(\prod_{j=3}^{i} \frac{n}{j[n+2(j-1)]}\right)\phi^i\right)}{\sum_{i=1}^{\infty}\left(\prod_{j=1}^{i} \frac{n}{j[n+2(j-1)]}\right)\phi^i}$$

Using that

$$\prod_{j=1}^{i} \frac{n}{j\left[\,n+2\,(j-1)\,\right]} = \frac{(n/2)^i}{i!}\prod_{j=1}^{i}\frac{1}{\left[\,\frac{n}{2}+(j-1)\,\right]} = \frac{(n/2)^i}{i!}\frac{\Gamma\left(\frac{n}{2}\right)}{\Gamma\left(\frac{n}{2}+i\right)}$$

we can write:

$$\begin{aligned}
\frac{Kur\left(\Delta p_i\right)}{N\left(\Delta p_i\right)} &= \frac{1}{\lambda}\frac{3n}{(n+2)} \frac{\left(\phi^2 + 2\,\Gamma\left(\frac{n}{2}+2\right)\,\left(\frac{2}{n}\right)^2 \sum_{i=3}^{\infty}\frac{1}{i!\,\Gamma\left(\frac{n}{2}+i\right)}\left(\phi\, n/2\right)^i\right)}{\Gamma\left(\frac{n}{2}\right)\,\sum_{i=1}^{\infty}\frac{1}{i!\,\Gamma\left(\frac{n}{2}+i\right)}\left(\phi\, n/2\right)^i} \\
&= \frac{1}{\lambda}\frac{3n}{(n+2)} \frac{2\,\Gamma\left(\frac{n}{2}+2\right)\,\left(\frac{2}{n}\right)^2}{\Gamma\left(\frac{n}{2}\right)} \frac{\left((1/2)\,\left(1/\Gamma\left(\frac{n}{2}+2\right)\right)\,\left(\frac{n}{2}\right)^2\phi^2 + \sum_{i=3}^{\infty}\frac{1}{i!\,\Gamma\left(\frac{n}{2}+i\right)}\left(\phi\, n/2\right)^i\right)}{\sum_{i=1}^{\infty}\frac{1}{i!\,\Gamma\left(\frac{n}{2}+i\right)}\left(\phi\, n/2\right)^i} \\
&= \frac{1}{\lambda}\frac{12n}{(n+2)} \frac{2\,\left(n/2+1\right)(n/2)}{n^2} \frac{\left(\frac{1}{2\Gamma\left(\frac{n}{2}+2\right)}\left(\phi\, n/2\right)^2 + \sum_{i=3}^{\infty}\frac{1}{i!\,\Gamma\left(\frac{n}{2}+i\right)}\left(\phi\, n/2\right)^i\right)}{\sum_{i=1}^{\infty}\frac{1}{i!\,\Gamma\left(\frac{n}{2}+i\right)}\left(\phi\, n/2\right)^i} \\
&= \frac{6}{\lambda}\frac{\sum_{i=2}^{\infty}\frac{1}{i!\,\Gamma\left(\frac{n}{2}+i\right)}\left(\phi\, n/2\right)^i}{\sum_{i=1}^{\infty}\frac{1}{i!\,\Gamma\left(\frac{n}{2}+i\right)}\left(\phi\, n/2\right)^i}
\end{aligned}$$

which gives the right hand side of equation (19). From there it is apparent that for a fixed

n, this ratio is increasing in $\lambda\bar{y}/\sigma^2$ and that for a fixed $\lambda\bar{y}/\sigma^2$, this ratio is increasing in n.

M Proof of Lemma 4

Proof. (of Lemma 4.) We use the property of the n independent BM's to write m as a function of a pair (z, y), where $z = \sum_i p_i$, as well as to write g as a function of (z, y) only. If each price gap follows an independent BM with common variance per unit of time σ^2, then, applying Ito's Lemma one can show that the pair (y, z) follows:

$$\begin{aligned} \mathrm{d}y(t) &= n\sigma^2 \mathrm{d}t + 2\sigma\sqrt{y(t)}\, \mathrm{d}\mathcal{W}^a(t) \\ \mathrm{d}z(t) &= \sqrt{n}\sigma \left[\frac{z(t)}{\sqrt{n\, y(t)}}\, \mathrm{d}\mathcal{W}^a(t) + \sqrt{1 - \left(\frac{z(t)}{\sqrt{n\, y(t)}}\right)^2}\, \mathrm{d}\mathcal{W}^b(t) \right] \end{aligned}$$

where $\mathcal{W}^a, \mathcal{W}^b$ are 2 standard (univariate) independent BM's. So that $\mathbb{E}(\mathrm{d}y)^2 = 4\sigma^2 y\, \mathrm{d}t$, $\mathbb{E}(\mathrm{d}z)^2 = n\sigma^2\, \mathrm{d}t$, and $\mathbb{E}(\mathrm{d}z\, \mathrm{d}y) = 2\sigma^2 z\, \mathrm{d}t$.

Hence we can write $\tilde{m}(p_1, ..., p_n) = \tilde{m}\left(||p||^2, \sum_{i=1}^n p_i\right)$, in which case $\tilde{m}$ solves the PDE :

$$\lambda\tilde{m}(z, y) = -z + \tilde{m}_y(z, y)n\sigma^2 + \tilde{m}_{zz}(z, y)\frac{n\sigma^2}{2} + \tilde{m}_{yy}(z, y)\frac{4\sigma^2 y}{2} + \tilde{m}_{zy}(z, y)\, 2\sigma^2 z$$

with boundary conditions: $\tilde{m}(z, \bar{y}) = 0$. We guess, and verify, that $\tilde{m}(z, y) = z\, \kappa_n(z)$ for some function $\kappa_n(\cdot)$ and where for emphasis we include the subindex n indicating the number of products. We then obtain:

$$\lambda\kappa_n(y) = -1 + \kappa_n'(y)\,(n+2)\sigma^2 + \kappa_n''(y)\, 2\, \sigma^2\, y$$

for all $0 \leq y \leq \bar{y}$ and $\kappa_n(\bar{y}) = 0$. Note that, except of the sign, this function obeys the same ODE and boundary conditions than the one for the time until adjustment $\mathcal{T}_{n+2}(y)$, which we solved to obtain $\mathcal{L}$ as if there were $n+2$ products instead of n products, and hence we get:

$$\kappa_n(y) = -\mathcal{T}_{n+2}(y) \tag{49}$$

The joint density of the invariant distribution $h(z, y)$ can be written as:

$$h(z, y) = s(z|y)\, f(y)$$

where f is the invariant distribution of y and $s(z|y)$ is the density distribution of the sum of the coordinates of a uniform distribution on an n dimensional hypersphere with square norm equal to y. In Alvarez and Lippi (2014) we have shown that this distribution is given by

$$s(z|y) = \frac{1}{Beta\left(\frac{n-1}{2},\frac{1}{2}\right)\ \sqrt{n\,y}} \left(1-\left(\frac{z}{\sqrt{y\,n}}\right)^2\right)^{(n-3)/2} \quad \text{for} \;\; z \in (-\sqrt{y\,n}\,,\ \sqrt{y\,n}) \tag{50}$$

Thus we can write

$$\mathcal{M}(\delta) \;\; = \;\; \frac{1}{\epsilon}\int_0^{\bar{y}}\int_{-\sqrt{ny}}^{\sqrt{ny}} \frac{1}{n}\,\tilde{m}(z-n\delta, y-2z\delta+n\delta^2)\ h(z,y)\ dz\ dy \tag{51}$$

and where we can express the invariant distribution of (z, y) with density h. Differentiating this expression w.r.t. δ and evaluating it at $\delta = 0$:

$$\begin{aligned}
\mathcal{M}'(0) \;\; &= \;\; -\frac{1}{n\epsilon}\int_0^{\bar{y}}\int_{-\sqrt{ny}}^{\sqrt{ny}} \left[n\,\frac{\partial \tilde{m}(z,y)}{\partial z} \;+\; \frac{\partial \tilde{m}(z,y)}{\partial y}\,2\,z\right]\ h(z,y)\ dz\ dy \\
&= \;\; -\frac{1}{n\epsilon}\left[\int_0^{\bar{y}} n\,\kappa_n(y)\,f(y)\,dy + 2\int_0^{\bar{y}}\kappa_n'(y)\int_{-\sqrt{ny}}^{\sqrt{ny}}\ z^2\ s(z|y)dz\,f(y)\ dy\right].
\end{aligned}$$

Integrating z^2 w.r.t. s gives $\int_{-\sqrt{ny}}^{\sqrt{ny}} z^2\ s(z|y)dz = y$ so

$$\begin{aligned}
\mathcal{M}'(0) \;\; &= \;\; -\frac{1}{n\epsilon}\int_0^{\bar{y}}\left[\ n\,\kappa_n(y) + 2\,\kappa_n'(y)y\ \right]f(y)\ dy \\
&= \;\; \frac{1}{\epsilon}\int_0^{\bar{y}}\left[\mathcal{T}_{n+2}(y) + \frac{2}{n}\,\mathcal{T}_{n+2}'(y)\,y\ \right]\ f(y)\ dy
\end{aligned}$$

where the last equality uses equation (49). □

N Power series representation of $\mathcal{T}_{n+2} + \mathcal{T}_{n+2}'\,y\,(2/n)$

Lemma 4 shows that $\partial m/\partial\delta$ can be written in terms of $\mathcal{T}_{n+2}$, the expected time until a price adjustment, as characterized in Proposition 3. In that proof we obtain the power series representation

$$\mathcal{T}_{n+2}(y) = \sum_{i=0}^{\infty}\alpha_{i,\,n+2}\,y^i$$

with

$$\alpha_{1,n+2} = \frac{1}{(\sigma^2/\lambda)(n+2)}\alpha_{0,n+2} - \frac{1}{\sigma^2(n+2)} = \frac{1}{(\sigma^2/\lambda)(n+2)}\left[\alpha_{0,n+2} - \frac{1}{\lambda}\right]$$

and for $i \geq 1$:

$$\alpha_{i+1,\, n+2} = \frac{\alpha_{i,\, n+2}}{(i+1)\,(\sigma^2/\lambda)\,(n+2+2i)} = \frac{\alpha_{i,\, n+2}}{(i+1)\,(\sigma^2/\lambda)\,(n/2+1+i)}\,\frac{1}{2}\,\left[\alpha_{0,n+2} - \frac{1}{\lambda}\right].$$

and using the properties of the Γ function:

$$\alpha_{i,\, n+2} = \frac{\Gamma\left(\frac{n}{2}+1\right)}{i!\,\Gamma\left(\frac{n}{2}+i+1\right)}\left(\frac{\lambda}{2\sigma^2}\right)^i \left[\alpha_{0,n+2} - \frac{1}{\lambda}\right]$$

Note that $\mathcal{T}_{n+2}(0) = \alpha_{0,\, n+2}$

Given the power series representation we have for all $y \in [0, \bar{y}]$:

$$\begin{aligned}
\mathcal{T}_{n+2}(y) + \mathcal{T}'_{n+2}(y)\, y\, \frac{2}{n} &= \sum_{i=0}^{\infty} \alpha_{i,\, n+2}\left[1 + i\frac{2}{n}\right] y^i \\
= \quad \alpha_{0,n+2} &+ \left[\alpha_{0,n+2} - \frac{1}{\lambda}\right] \sum_{i=1}^{\infty} \frac{\Gamma\left(\frac{n}{2}+1\right)}{i!\,\Gamma\left(\frac{n}{2}+i+1\right)}\left[1 + i\frac{2}{n}\right]\left(\frac{\lambda y}{2\sigma^2}\right)^i
\end{aligned}$$

Note that $\alpha_{0,n+2} = \mathcal{T}_{n+2}(0)$ with

$$\begin{aligned}
\lambda\,\alpha_{0,n+2} = \ell &= \frac{\sum_{i=1}^{\infty}\left(\prod_{j=1}^{i}\frac{1}{j[n+2+2(j-1)]}\right)\left(\frac{\lambda\bar{y}}{\sigma^2}\right)^i}{1 + \sum_{i=1}^{\infty}\left(\prod_{j=1}^{i}\frac{1}{j[n+2+2(j-1)]}\right)\left(\frac{\lambda\bar{y}}{\sigma^2}\right)^i} = \frac{\sum_{i=1}^{\infty}\left(\prod_{j=1}^{i}\frac{1}{j\left[\frac{n}{2}+j\right]}\right)\left(\frac{\lambda\bar{y}}{2\,\sigma^2}\right)^i}{1 + \sum_{i=1}^{\infty}\left(\prod_{j=1}^{i}\frac{1}{j\left[\frac{n}{2}+j\right]}\right)\left(\frac{\lambda\bar{y}}{2\,\sigma^2}\right)^i} \\
&= \frac{\sum_{i=1}^{\infty}\frac{\Gamma\left(\frac{n}{2}+1\right)}{i!\,\Gamma\left(\frac{n}{2}+1+i\right)}\left(\frac{\lambda\bar{y}}{2\,\sigma^2}\right)^i}{\sum_{i=0}^{\infty}\frac{\Gamma\left(\frac{n}{2}+1\right)}{i!\,\Gamma\left(\frac{n}{2}+1+i\right)}\left(\frac{\lambda\bar{y}}{2\,\sigma^2}\right)^i}
\end{aligned}$$

Thus we have:

$$\lambda\left(\mathcal{T}_{n+2}(y)+\mathcal{T}'_{n+2}(y)\,y\,\frac{2}{n}\right)=\frac{\sum_{i=1}^{\infty}\frac{\Gamma\left(\frac{n}{2}+1\right)}{i!\,\Gamma\left(\frac{n}{2}+i+1\right)}\left(\frac{\lambda\bar{y}}{2\,\sigma^2}\right)^i}{\sum_{i=0}^{\infty}\frac{\Gamma\left(\frac{n}{2}+1\right)}{i!\,\Gamma\left(\frac{n}{2}+i+1\right)}\left(\frac{\lambda\bar{y}}{2\,\sigma^2}\right)^i}$$

$$+\left[\frac{\sum_{i=1}^{\infty}\frac{\Gamma\left(\frac{n}{2}+1\right)}{i!\,\Gamma\left(\frac{n}{2}+i+1\right)}\left(\frac{\lambda\bar{y}}{2\,\sigma^2}\right)^i}{\sum_{i=0}^{\infty}\frac{\Gamma\left(\frac{n}{2}+1\right)}{i!\,\Gamma\left(\frac{n}{2}+i+1\right)}\left(\frac{\lambda\bar{y}}{2\,\sigma^2}\right)^i}-1\right]\left[\sum_{i=1}^{\infty}\left[1+i\frac{2}{n}\right]\frac{\Gamma\left(\frac{n}{2}+1\right)}{i!\,\Gamma\left(\frac{n}{2}+i+1\right)}\left(\frac{\lambda y}{2\sigma^2}\right)^i\right]$$

$$=\frac{\sum_{i=1}^{\infty}\frac{\Gamma\left(\frac{n}{2}+1\right)}{i!\,\Gamma\left(\frac{n}{2}+i+1\right)}\left(\frac{1}{2}\right)^i\left(\frac{\lambda\bar{y}}{\sigma^2}\right)^i-\sum_{i=1}^{\infty}\frac{\Gamma\left(\frac{n}{2}+1\right)}{i!\,\Gamma\left(\frac{n}{2}+i+1\right)}\left[1+i\frac{2}{n}\right]\left(\frac{\lambda y}{2\,\sigma^2}\right)^i}{\sum_{i=0}^{\infty}\frac{\Gamma\left(\frac{n}{2}+1\right)}{i!\,\Gamma\left(\frac{n}{2}+i+1\right)}\left(\frac{1}{2}\right)^i\left(\frac{\lambda\bar{y}}{\sigma^2}\right)^i}$$

$$=\frac{\sum_{i=1}^{\infty}\frac{\Gamma\left(\frac{n}{2}+1\right)}{i!\,\Gamma\left(\frac{n}{2}+i+1\right)}\left[\left(\frac{\lambda\bar{y}}{2\,\sigma^2}\right)^i-\left(1+\frac{2i}{n}\right)\left(\frac{\lambda y}{2\sigma^2}\right)^i\right]}{\sum_{i=0}^{\infty}\frac{\Gamma\left(\frac{n}{2}+1\right)}{i!\,\Gamma\left(\frac{n}{2}+i+1\right)}\left(\frac{\lambda\bar{y}}{2\,\sigma^2}\right)^i} \tag{52}$$

We can write this as:

$$\lambda\left(\mathcal{T}_{n+2}(y)+\mathcal{T}'_{n+2}(y)\,y\,\frac{2}{n}\right)=\frac{\sum_{i=1}^{\infty}\gamma_i}{\sum_{i=0}^{\infty}\gamma_i}-\frac{\sum_{i=1}^{\infty}\gamma_i\left(1+\frac{2\,i}{n}\right)\left(\frac{y}{\bar{y}}\right)^i}{\sum_{i=0}^{\infty}\gamma_i} \tag{53}$$

where

$$\gamma_i=\frac{\Gamma\left(\frac{n}{2}+1\right)}{i!\,\Gamma\left(\frac{n}{2}+1+i\right)}\left(\frac{\lambda\bar{y}}{2\sigma^2}\right)^i \tag{54}$$

O Power series representation of the density $f(y)$

From equation (7) we can write f as the product of a power of y and the sums of two modified Bessel functions of the first and second kind, multiplied by appropriate constants.

Consider then $n\geq 3$ and n odd, so that $\nu=n/2-1$ is not an integer. When n is even the expression for K_ν requires to evaluate the limit, so it is more complicated. Thus, we can write:

$$\left(\frac{\lambda y}{2\sigma^2}\right)^{\left(\frac{n}{4}-\frac{1}{2}\right)} I_{\frac{n}{2}-1}\left(2\sqrt{\frac{\lambda y}{2\sigma^2}}\right)\;=\;\left(\frac{\lambda y}{2\sigma^2}\right)^{\left(\frac{n}{2}-1\right)}\sum_{i=0}^{\infty}\beta_{i,\frac{n}{2}-1}\left(\frac{\lambda y}{2\sigma^2}\right)^i$$

where

$$\beta_{i,\frac{n}{2}-1}\equiv\frac{1}{i!\,\Gamma(i+n/2)}$$

and for ν not an integer

$$\begin{aligned}
\left(\frac{\lambda y}{2\sigma^2}\right)^{\left(\frac{n}{4}-\frac{1}{2}\right)} K_{\frac{n}{2}-1}\left(2\sqrt{\frac{\lambda y}{2\sigma^2}}\right) &= \frac{\frac{\pi}{2}}{\sin\left(\left(\frac{n}{2}-1\right)\pi\right)} \sum_{i=0}^{\infty} \beta_{i,1-\frac{n}{2}} \left(\frac{\lambda y}{2\sigma^2}\right)^i \\
&- \frac{\frac{\pi}{2}}{\sin\left(\left(\frac{n}{2}-1\right)\pi\right)} \left(\frac{\lambda y}{2\sigma^2}\right)^{\left(\frac{n}{4}-\frac{1}{2}\right)} I_{\frac{n}{2}-1}\left(2\sqrt{\frac{\lambda y}{2\sigma^2}}\right)
\end{aligned}$$

where

$$\beta_{i,1-\frac{n}{2}} \equiv \frac{1}{i!\,\Gamma(i+2-n/2)}$$

This means we can write:

$$\begin{aligned}
f(y) &= \left(C_I - \frac{\frac{\pi}{2}}{\sin\left(\left(\frac{n}{2}-1\right)\pi\right)} C_K\right) \left(\frac{\lambda y}{2\sigma^2}\right)^{\left(\frac{n}{2}-1\right)} \sum_{i=0}^{\infty} \beta_{i,\frac{n}{2}-1} \left(\frac{\lambda y}{2\sigma^2}\right)^i \\
&+ C_K \frac{\frac{\pi}{2}}{\sin\left(\left(\frac{n}{2}-1\right)\pi\right)} \sum_{i=0}^{\infty} \beta_{i,1-\frac{n}{2}} \left(\frac{\lambda y}{2\sigma^2}\right)^i
\end{aligned}$$

Since $f(0) > 0$ and

$$f(0) = C_K \frac{\frac{\pi}{2}}{\sin\left(\left(\frac{n}{2}-1\right)\pi\right)} \beta_{0,1-\frac{n}{2}} = C_K \frac{\frac{\pi}{2}}{\sin\left(\left(\frac{n}{2}-1\right)\pi\right)} \frac{1}{\Gamma(2-n/2)}$$

then $C_K > 0$. Then to set $f(\bar{y}) = 0$ we obtain:

$$\frac{C_I - \frac{\frac{\pi}{2}}{\sin\left(\left(\frac{n}{2}-1\right)\pi\right)} C_K}{-\frac{\frac{\pi}{2}}{\sin\left(\left(\frac{n}{2}-1\right)\pi\right)} C_K} = \frac{\sum_{i=0}^{\infty} \beta_{i,1-\frac{n}{2}} \left(\frac{\lambda \bar{y}}{2\sigma^2}\right)^i}{\left(\frac{\lambda \bar{y}}{2\sigma^2}\right)^{\left(\frac{n}{2}-1\right)} \sum_{i=0}^{\infty} \beta_{i,\frac{n}{2}-1} \left(\frac{\lambda \bar{y}}{2\sigma^2}\right)^i}$$

Using the expressions for $f(0)$ and $f(\bar{y}) = 0$ we can then rewrite f as:

$$\begin{aligned}
f(y) &= -f(0)\,\Gamma\left(2-\frac{n}{2}\right) \left(\sum_{i=0}^{\infty} \beta_{i,1-\frac{n}{2}} \left(\frac{\lambda \bar{y}}{2\sigma^2}\right)^i\right) \times \\
&\left[\frac{\left(\frac{\lambda y}{2\sigma^2}\right)^{\left(\frac{n}{2}-1\right)} \sum_{i=0}^{\infty} \beta_{i,\frac{n}{2}-1} \left(\frac{\lambda y}{2\sigma^2}\right)^i}{\left(\frac{\lambda \bar{y}}{2\sigma^2}\right)^{\left(\frac{n}{2}-1\right)} \sum_{i=0}^{\infty} \beta_{i,\frac{n}{2}-1} \left(\frac{\lambda \bar{y}}{2\sigma^2}\right)^i} - \frac{\sum_{i=0}^{\infty} \beta_{i,1-\frac{n}{2}} \left(\frac{\lambda y}{2\sigma^2}\right)^i}{\sum_{i=0}^{\infty} \beta_{i,1-\frac{n}{2}} \left(\frac{\lambda \bar{y}}{2\sigma^2}\right)^i}\right]
\end{aligned}$$

Using that $1 = \int_0^{\bar{y}} f(y)dy$ we obtain an expression for $f(0)$ and replacing in the previous formula we obtain:

$$f(y) = \left[\frac{\left(\frac{\lambda y}{2\sigma^2}\right)^{\left(\frac{n}{2}-1\right)} \sum_{i=0}^{\infty} \beta_{i,\frac{n}{2}-1} \left(\frac{\lambda y}{2\sigma^2}\right)^i}{\left(\frac{\lambda \bar{y}}{2\sigma^2}\right)^{\left(\frac{n}{2}-1\right)} \sum_{i=0}^{\infty} \beta_{i,\frac{n}{2}-1} \left(\frac{\lambda \bar{y}}{2\sigma^2}\right)^i} - \frac{\sum_{i=0}^{\infty} \beta_{i,1-\frac{n}{2}} \left(\frac{\lambda y}{2\sigma^2}\right)^i}{\sum_{i=0}^{\infty} \beta_{i,1-\frac{n}{2}} \left(\frac{\lambda \bar{y}}{2\sigma^2}\right)^i} \right] \Big/ \left[\frac{\left(\frac{\lambda \bar{y}}{2\sigma^2}\right)^{\left(\frac{n}{2}-1\right)} \sum_{i=0}^{\infty} \beta_{i,\frac{n}{2}-1} \frac{\bar{y}}{i+n/2} \left(\frac{\lambda \bar{y}}{2\sigma^2}\right)^i}{\left(\frac{\lambda \bar{y}}{2\sigma^2}\right)^{\left(\frac{n}{2}-1\right)} \sum_{i=0}^{\infty} \beta_{i,\frac{n}{2}-1} \left(\frac{\lambda \bar{y}}{2\sigma^2}\right)^i} - \frac{\sum_{i=0}^{\infty} \beta_{i,1-\frac{n}{2}} \frac{\bar{y}}{i+1} \left(\frac{\lambda \bar{y}}{2\sigma^2}\right)^i}{\sum_{i=0}^{\infty} \beta_{i,1-\frac{n}{2}} \left(\frac{\lambda \bar{y}}{2\sigma^2}\right)^i} \right] \tag{55}$$

Remark. While this expression was obtained for $n \geq 2$ and n odd, it does work for any real number $n \geq 2$ different from an even natural. Since it is continuous in n, the expression equation (55) can be used to obtain the values of f in the case of n is even by taking the limit as n approaches any even natural, or by evaluating at a real number very close to the desired even natural number.

P Discrete Time Formulation for Proposition 12.

We start with discrete time version of the process for price gaps, with length of the time period Δ, which makes some of the arguments more accessible. Let N be

$$N(t+\Delta) = \begin{cases} N(t) & \text{with probability } (1-\lambda\Delta) \\ N(t)+1 & \text{with probability } \lambda\Delta \end{cases} \tag{56}$$

Thus, as $\Delta \downarrow 0$ this process converges to a continuous time Poisson counter with instantaneous intensity rate λ per unit of time. Let $\bar{p}_i$ follow n drift-less random walks

$$\bar{p}_i(t+\Delta, p) = \begin{cases} \bar{p}_i(t,p) + \sigma\sqrt{\Delta} & \text{with probability } 1/2 \\ \bar{p}_i(t,p) - \sigma\sqrt{\Delta} & \text{with probability } 1/2 \end{cases} \tag{57}$$

where the initial condition satisfies:

$$\bar{p}_i(0) = p_i \text{ for } i = 1, .., n \ ,$$

and where the n random walks are independent of each other and of the Poisson counter. As $\Delta \downarrow 0$ the process for $\bar{p}$ converges to a Brownian motion whose changes have variance σ^2 per unit of time. We define the stopping time of the first price adjustment $\tau(p)$, conditional on the starting at price gap vector p at time zero, as:

$$\begin{aligned}
\tau_1 &\equiv \min\left\{t = 0, \Delta, 2\Delta, ... : N(j\Delta + \Delta) - N(j\Delta) = 1\right\}, \\
\tau_2(p) &\equiv \min\left\{t = 0, \Delta, 2\Delta, ... : \sum_{i=1}^{n} \left(\bar{p}_i(j\Delta + \Delta, p)\right)^2 \geq \bar{y}\right\} \text{ and} \\
\tau(p) &\equiv \min\left\{\tau_1 \, , \, \tau_2(p)\right\}.
\end{aligned}$$

The function g is the density for the continuous time limit, i.e. the case where $\Delta \downarrow 0$. For small Δ, we can approximate the distribution of the *fraction* of firms with price gap vector p as the product of the density g and a correction to convert it into a probability, i.e a fraction. This gives:

$$g\left(p_1, ..., p_; n, \lambda/\sigma^2, \bar{y}\right) \, \left(\sigma\sqrt{\Delta}\right)^n$$

where the last term uses that in each dimension price gaps vary discretely in steps of size $\sigma\sqrt{\Delta}$. We can write the discrete time impulse response function as:

$$\mathcal{P}(t, \delta; \sigma, \lambda, \bar{y}, \Delta) \;=\; \Theta(\delta; \sigma, \lambda, \bar{y}, \Delta) + \sum_{s=\Delta}^{t} \theta(\delta, s; \sigma, \lambda, \bar{y}, , \Delta)\,\Delta\,,$$

In this expression we can, without loss of generality, restrict t to be an integer multiple of Δ. We have divided the expression for θ by Δ, and hence multiplied its contribution back by Δ in $\mathcal{P}$, so that it has the interpretation of the contribution per unit of time to the IRF of price changes at time t, i.e. it has the units of a density. Moreover, in this manner the term has a non-zero limit, and the expression in $\mathcal{P}$ converges to an integral. Thus we get the $\mathcal{P} = \lim \mathcal{P}(\Delta)$ as $\Delta \downarrow \infty$. The functions θ and Θ are given by:

$$\Theta(\delta; \sigma, \lambda, \bar{y}, \Delta) \equiv \sum_{||p(0) - \iota\delta|| \geq \bar{y}} \left(\delta - \frac{\sum_{j=0}^{n} p_j(0)}{n}\right) g\left(p(0); n, \frac{\lambda}{\sigma^2}, \bar{y}\right) \left(\sigma\sqrt{\Delta}\right)^n \text{, and}$$

$$\begin{aligned}
&\theta(\delta, t; \sigma, \lambda, \bar{y}, \Delta) \equiv \\
&-\frac{1}{\Delta} \sum_{||p(0) - \iota\delta|| < \bar{y}} \mathbb{E}\left[\frac{\sum_{j=0}^{n} \bar{p}_j(t, p)}{n} \, \mathbf{1}_{\{\tau(p) = t\}} \,\middle|\, p = p(0) - \iota\delta\right] g\left(p(0); n, \frac{\lambda}{\sigma^2}, \bar{y}\right) \left(\sigma\sqrt{\Delta}\right)^n
\end{aligned}$$

Time scaling of the IRF with $N(\Delta p_i)$. For this (i) Note that if multiply the parameters σ^2 and λ by a constant $k > 0$, leaving $\bar{y}$ unaltered, then $N(\Delta p_i)' = k\, N(\Delta p_i)$, where primes are used to denote the values that correspond to the scaled parameters. This follows directly from the expression we derive for $N(\Delta p_i) = 1/T(0)$ in Proposition 3. (ii) By Proposition 4 with these changes the distribution of price changes implied by $(\sigma^2, \lambda, \bar{y})$ is exactly the same as the one implied by $(k\sigma^2, k\lambda, \bar{y})$. (iii) we change notation and write $(\sigma^2, \lambda, \bar{y})$ instead of $(\lambda, \sigma^2, \psi/B)$ and omit n. We establish that

$$\mathcal{P}_n\left(\frac{t}{k}, \delta; k\sigma^2, k\lambda, \bar{y}\right) = \mathcal{P}_n\left(t, \delta; \sigma^2, \lambda, \bar{y}\right)$$

We will do so by establishing this proposition for the discrete time version of the IRF. Yet the result is immediate, since λ and σ^2 are the only two parameters which are rates per unit of time (the other parameters are n and $\bar{y}$), so by multiplying them by k we just scale time. The details can be found in the discrete time formulation, whose notation we develop below. We show that

$$\mathcal{P}(t, \delta; k\sigma^2, k\lambda, \bar{y}, \Delta/k) = \mathcal{P}(t/k, \delta; \sigma^2, \lambda, \bar{y}, \Delta) \tag{58}$$

We will do so by establishing this proposition for the discrete time version of the IRF. Let $\Delta' = \Delta/k$, $\sigma'^2 = \sigma^2 k$ and $\lambda' = \lambda k$. Note that, by construction $\sigma' \sqrt{\Delta'} = \sigma \sqrt{\Delta}$ and $\lambda'/(\sigma')^2 = \lambda/(\sigma)^2$. To establish this we first note that, for a given shock δ, Θ depends only on n, $\bar{y}$, $\sigma\sqrt{\Delta}$, and λ/σ^2. This is because the invariant density g and the scaling factor to convert it into probabilities depends only on those parameters. Second we show that

$$\sum_{s=\Delta/k}^{t/k} \frac{\Delta}{k} \theta\left(s, \delta; k\sigma^2, k\lambda, \bar{y}, \frac{\Delta}{k}\right) = \sum_{s=\Delta}^{t} \Delta\, \theta\left(s, \delta; \sigma, \lambda, \bar{y}, \Delta\right)$$

This follows because for each s and $p(0)$

$$\begin{aligned} & \mathbb{E}\left[\frac{\sum_{j=0}^{n} \bar{p}_j\left(s, p\right)}{n} \; \mathbf{1}_{\{\tau(p)=s\}} \;\middle|\; p = p(0) - \iota\delta\,;\, \sigma, \lambda, \Delta\right] \\ = \; & \mathbb{E}\left[\frac{\sum_{j=0}^{n} \bar{p}_j\left(\frac{s}{k}, p\right)}{n} \; \mathbf{1}_{\left\{\tau(p)=\frac{s}{k}\right\}} \;\middle|\; p = p(0) - \iota\delta\,;\, \sigma', \lambda', \Delta'\right] \end{aligned}$$

where we include the parameters $(\lambda, \sigma^2, \Delta)$ as argument of the expected values. This itself follows because, using equation (56) and equation (57) then the processes for $\{\bar{p}_i\}$ are the same in the original time and in the time time scales by k since the probabilities of the counter to go up $\lambda'\Delta' = \lambda\Delta$ and the steps of the symmetric random walks $\sigma'\sqrt{\Delta'} = \sigma\sqrt{\Delta}$ are the same in the original time and the time scaled by k. In particular we have that

$$\bar{p}_j\left(\frac{s}{k}, p; \lambda', \sigma'^2, \Delta'\right) \equiv \bar{p}_j\left(\frac{s}{k}, p; k\lambda, k\sigma^2, \frac{\Delta}{k}\right) = \bar{p}_j\left(s, p; \lambda, \sigma^2, \Delta\right) = \hat{p}$$

with exactly the same probabilities for each price gap $\hat{p} \in \mathbb{R}$ and each time $s \geq 0$. Also, repeating the arguments used for Θ, we have $g\left(p(0); n, \frac{\lambda}{\sigma^2}, \bar{y}\right) \left(\sigma\sqrt{\Delta}\right)^n = g\left(p(0); n, \frac{\lambda'}{\sigma'^2}, \bar{y}\right) \left(\sigma'\sqrt{\Delta'}\right)^n$. Thus, since equation (58) holds for all $\Delta > 0$, taking limits

$$\mathcal{P}\left(\frac{t}{k}, \delta; k\sigma^2, k\lambda, \bar{y}\right) = \lim_{\Delta\downarrow 0} \mathcal{P}\left(\frac{t}{k}, \delta; k\sigma^2, k\lambda, \bar{y}, \frac{\Delta}{k}\right) = \lim_{\Delta\downarrow 0} \mathcal{P}\left(t, \delta; \sigma^2, \lambda, \bar{y}, \Delta\right) = \mathcal{P}\left(t, \delta; \sigma^2, \lambda, \bar{y}\right)$$

Scaling of the IRF in the monetary shock with $Std(\Delta p_i)$. For this we use properties of the invariant distribution f, which are then inherited by g. In particular, we will compare the IRF with parameters $(\lambda, \sigma^2, \bar{y})$ with one with parameters $(\lambda', \sigma'^2, \bar{y})$ where $\lambda' = \lambda$, $\sigma'^2 = k\,\sigma^2$ and $\bar{y}' = k\,\bar{y}$. With this choice we have $N(\Delta p_i)' = N(\Delta p_i)$ and thus $\ell = \lambda'/N(\Delta p_i)'$ since $\lambda\bar{y}/(n\sigma^2) = \lambda'\bar{y}'/(n\sigma'^2)$ (see Proposition 3). Then by Proposition 1 we have that the standard deviation of price changes scales up with k, i.e.: $Std(\Delta p_i)' = \sqrt{k}\,Std(\Delta p_i)$. The main idea is that the invariant distribution corresponding to the $'$ parameters is a radial expansion of the original, so that $\int_0^y f(x; \lambda, \sigma^2, \bar{y})dx = \int_0^{yk} f(x; \lambda', \sigma'^2, \bar{y}')dx$ and thus $f(y, \lambda, \sigma^2, \bar{y}) = kf(yk, \lambda', \sigma'^2, \bar{y}')$. Indeed using Lemma 3 we have:

$$f\left(y; \frac{\lambda}{\sigma^2}, \bar{y}\right) = k\,f\left(yk; \frac{\lambda}{k\sigma^2}, k\bar{y}\right) \equiv k\,f\left(yk; \frac{\lambda'}{\sigma'^2}, \bar{y}'\right) \quad . \tag{59}$$

Thus we have:

$$\begin{aligned}
g\left(p_1, ..., p_; n, \frac{\lambda}{\sigma^2}, \bar{y}\right) &= f\left(p_1^2 + \cdots + p_n^2; n, \frac{\lambda}{\sigma^2}, \bar{y}\right) \frac{\Gamma(n/2)}{2\,\pi^{n/2}\ \left(p_1^2 + \cdots + p_n^2\right)^{(n-2)/2}} = \\
&= kf\left(k(p_1^2 + \cdots + p_n^2); n, \frac{\lambda'}{\sigma'^2}, \bar{y}'\right) \frac{\Gamma(n/2)\,k^{(n-1)/2}}{2\,\pi^{n/2}\ \left(k(p_1^2 + \cdots + p_n^2)\right)^{(n-2)/2}} \\
&= g\left(\sqrt{k}(p_1, ..., p_n); n, \frac{\lambda'}{\sigma'^2}, \bar{y}'\right) k^{(n-2)/2}\,k
\end{aligned}$$

Using this for the discrete time formulation we have:

$$\begin{aligned}
g\left(p; n, \frac{\lambda}{\sigma^2}, \bar{y}\right) \left(\sigma\sqrt{\Delta}\right)^n &= g\left(\sqrt{k}\,p; n, \frac{\lambda'}{\sigma'^2}, \bar{y}'\right) \left(\sigma'\sqrt{\Delta}\right)^n k^{(n-2)/2}\,k\,k^{-n/2} \\
&= g\left(\sqrt{k}\,p; n, \frac{\lambda'}{\sigma'^2}, \bar{y}'\right) \left(\sigma'\sqrt{\Delta}\right)^n
\end{aligned}$$

Note that $\{||p(0) - \iota\delta|| \geq \bar{y}\} = \{||\sqrt{k}\,p(0) - \iota\sqrt{k}\,\delta|| \geq \sqrt{k}\,\bar{y}\} = \{||\sqrt{k}\,p(0) - \iota\delta'|| \geq \bar{y}'\}$. Also

$$\left(\delta - \frac{\sum_{j=0}^n p_j(0)}{n}\right)\sqrt{k} = \left(\delta' - \frac{\sum_{j=0}^n \sqrt{k}\,p_j(0)}{n}\right)$$

Thus

$$\sqrt{k} \sum_{||p(0)-\iota\delta||\geq\bar{y}} \left(\delta - \frac{\sum_{j=0}^{n} p_j(0)}{n}\right) g\left(p(0); n, \frac{\lambda}{\sigma^2}, \bar{y}\right) \left(\sigma\sqrt{\Delta}\right)^n$$
$$= \sum_{||\sqrt{k}p(0)-\iota\delta'||\geq\bar{y}'} \left(\delta' - \frac{\sum_{j=0}^{n} \sqrt{k}p_j(0)}{n}\right) g\left(\sqrt{k}\, p(0); n, \frac{\lambda'}{\sigma'^2}, \bar{y}'\right) \left(\sigma'\sqrt{\Delta}\right)^n$$

Using the definition of $\Theta(\cdot, \Delta)$:

$$\sqrt{k}\, \Theta(\delta; \sigma, \lambda, \bar{y}, \Delta) = \Theta\left(\sqrt{k}\,\delta; k\sigma^2, \lambda, k\bar{y}, \Delta\right) \equiv \Theta\left(\delta'; \sigma'^2, \lambda', \bar{y}'\Delta\right) .$$

Since this holds for all Δ, by taking limits as $\Delta \downarrow 0$, we have shown the desired result for Θ. The result for θ follows the steps for g. We set $\Delta' = \Delta$ and note that for all $p(0) \in \mathbb{R}^n$, scaling factor $k > 0$ and time horizon $s > 0$:

$$\sqrt{k}\; \mathbb{E}\left[\frac{\sum_{j=0}^{n} \bar{p}_j\left(s, p\right)}{n}\; \mathbf{1}_{\{\tau(p)=s\}} \;\middle|\; p = p(0) - \iota\delta\, ;\, \sigma, \lambda, \Delta\right]$$
$$= \mathbb{E}\left[\frac{\sum_{j=0}^{n} \bar{p}_j\left(s, p\right)}{n}\; \mathbf{1}_{\{\tau(p)=s\}} \;\middle|\; p = \sqrt{k}\, p(0) - \iota\delta'\, ;\, \sigma', \lambda', \Delta\right] .$$

This follows because $\lambda' = \lambda$ and $\sigma'\sqrt{\Delta'} = \sqrt{k}\,\sigma\,\sqrt{\Delta}$, thus the each $p \in \mathbb{R}^n$ the paths $\sqrt{k}\,\bar{p}(s, p; \sigma, \lambda) = \bar{p}(s, \sqrt{k}p; \sigma', \lambda')$ occur with the same probabilities.

Q Detailed Proof. of Proposition 8.

Proof. (of Proposition 8.) In general we have $\underline{\delta} = 2\sqrt{\bar{y}/n}$, since for a shock of this size every single firm for which $||p||^2 = y \leq \bar{y}$ before the shock will find that $||p - \iota\delta||^2 \geq \bar{y}$, where ι is a vector of ones. In particular we want to find out the smallest value of δ for which

$$||p - \iota\delta||^2 = ||p||^2 - 2\,\delta \sum_i p_i + n\delta^2 \geq \bar{y}$$

for any $||p||^2 \leq \bar{y}$. Using that $\sum_i p_i \leq n\sqrt{y/n}$ for $y = ||p||^2$ it is easy to establish the desired result.

We can rewrite it as $\underline{\delta} = 2\sqrt{\bar{y}/n} = 2\sqrt{\sigma^2/\lambda}\,\sqrt{\phi}$, which gives an equivalent way to write the expression for $\underline{\delta}$ as

$$\underline{\delta} = Std(\Delta p_i)\, 2 \sqrt{\frac{\phi}{\mathcal{L}(\phi, n)}} \text{ where } \phi \equiv \bar{y}\, \lambda/(n\sigma^2) \ .$$

where $\phi(n, \ell) \equiv \bar{y}\lambda/(n\sigma^2)$ a function that depends only on ℓ and n, as shown in Proposition 3. Using Proposition 1 we have:

$$(N(\Delta p_i)/\lambda)\ Var(\Delta p_i) = \sigma^2/\lambda \text{ or } \sigma^2/\lambda = Var(\Delta p_i)/\ell$$

Combining the two equations we obtain the desired result.

Note that $\phi(\ell, n)/\ell = \phi/\mathcal{L}(\phi, n)$. Since $\mathcal{L}(\phi, n)$ is increasing in ϕ with $\lim_{\phi\to\infty} \mathcal{L}(\phi, n) = 1$, then $\lim_{\ell\to 1} \phi(\ell, n)/\ell = \infty$. To study the limit as $\ell \to 0$, using the functional form of $\mathcal{L}$, and taking a Taylor expansion of $\mathcal{L}(\phi, n) = \phi + o(\phi)$, thus

$$\frac{\phi}{\mathcal{L}(\phi, n)} = \frac{\phi}{\phi + o(\phi)} = \frac{1}{1 + o(\phi)/\phi} \, ,$$

and hence

$$\lim_{\ell\to 0} \frac{\phi(\ell, n)}{\ell} = \lim_{\phi\to 0} \frac{\phi}{\mathcal{L}(\phi, n)} = 1 \, .$$

Omitting n to simplify the notation we have:

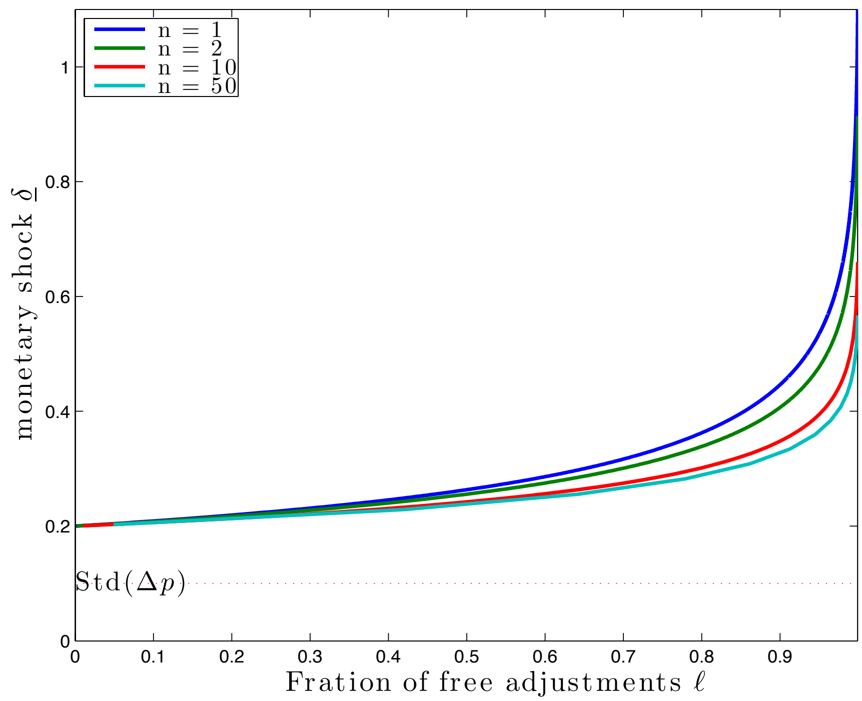
$$\frac{\partial}{\partial \phi}\left[\frac{\phi}{\mathcal{L}(\phi)}\right] = \frac{1}{\mathcal{L}(\phi)}\left[1 - \frac{\mathcal{L}'(\phi)\phi}{\mathcal{L}(\phi)}\right]$$

and rewriting $\mathcal{L}(\phi) = \frac{g(\phi)}{1+g(\phi)}$ we obtain: $\mathcal{L}'(\phi) = \frac{g'(\phi)}{[1+g(\phi)]^2}$ and thus

$$\frac{\mathcal{L}'(\phi)\phi}{\mathcal{L}(\phi)} = \frac{g'(\phi)}{(1 + g(\phi))^2} \frac{(1 + g(\phi))}{g(\phi)} \phi = \frac{g'(\phi)}{(1 + g(\phi))} \frac{\phi}{g(\phi)}$$

since $g(\cdot)$ is convex and $g(0) = 0$ then $0 = g(0) \geq g(\phi) + g'(\phi)(0 - \phi)$ or $g(\phi) \leq g'(\phi)\phi$ $\frac{\mathcal{L}'(\phi)\phi}{\mathcal{L}(\phi)} \leq \frac{1}{1+g(\phi)} \leq 1$ and thus $\phi(\ell, n)/\ell$ is strictly increasing in ℓ for all $\ell \in (0, 1)$. $\square$

Figure 7: Minimum size of monetary shock for full price flexibility



R Proof of Lemma 5

Proof. (of Lemma 5) We can rewrite this expression as

$$\frac{\lambda\, Kur\,(\Delta p_i)}{6\, N\,(\Delta p_i)} = \frac{\sum_{i=1}^{\infty} \frac{\Gamma\left(\frac{n}{2}+1\right)}{i!\,\Gamma\left(\frac{n}{2}+1+i\right)} \left(\frac{\lambda \bar{y}}{2\sigma^2}\right)^i \frac{1}{1+i}}{\sum_{i=0}^{\infty} \frac{\Gamma\left(\frac{n}{2}+1\right)}{i!\,\Gamma\left(\frac{n}{2}+1+i\right)} \left(\frac{\lambda \bar{y}}{2\sigma^2}\right)^i \frac{1}{1+i}} = \frac{\sum_{i=1}^{\infty} \gamma_i \frac{1}{1+i}}{\sum_{i=0}^{\infty} \gamma_i \frac{1}{1+i}} \tag{60}$$

Thus the equation

$$\frac{\lambda\, Kur\,(\Delta p_i)}{6\, N\,(\Delta p_i)} = \int_0^{\bar{y}} \left[\lambda \left(\mathcal{T}_{n+2}(y) + \mathcal{T}'_{n+2}(y)\, y\, \frac{2}{n}\right)\right] f(y)\, dy \tag{61}$$

is equivalent to:

$$\frac{\sum_{i=1}^{\infty} \gamma_i \frac{1}{1+i}}{\sum_{i=0}^{\infty} \gamma_i \frac{1}{1+i}} - \frac{\sum_{i=1}^{\infty} \gamma_i}{\sum_{i=0}^{\infty} \gamma_i} = -\frac{\sum_{i=1}^{\infty} \gamma_i \left(1 + \frac{2\,i}{n}\right)}{\sum_{i=0}^{\infty} \gamma_i} \int_0^{\bar{y}} \left(\frac{y}{\bar{y}}\right)^i f(y)\, dy$$

We can write this equation as:

$$\frac{\left(\sum_{i=0}^{\infty}\gamma_i\right)\left(\sum_{i=1}^{\infty}\gamma_i\,\frac{1}{1+i}\right)-\left(\sum_{i=1}^{\infty}\gamma_i\right)\left(\sum_{i=0}^{\infty}\gamma_i\,\frac{1}{1+i}\right)}{\left(\sum_{i=0}^{\infty}\gamma_i\,\frac{1}{1+i}\right)\left(\sum_{i=0}^{\infty}\gamma_i\right)}$$
$$=-\frac{\sum_{i=1}^{\infty}\gamma_i\,\left(1+\frac{2\,i}{n}\right)}{\sum_{i=0}^{\infty}\gamma_i}\int_0^{\bar{y}}\left(\frac{y}{\bar{y}}\right)^i f(y)\,dy$$

or

$$\frac{\left(\gamma_0+\sum_{i=1}^{\infty}\gamma_i\right)\left(\sum_{i=1}^{\infty}\gamma_i\,\frac{1}{1+i}\right)-\left(\sum_{i=1}^{\infty}\gamma_i\right)\left(\gamma_0+\sum_{i=1}^{\infty}\gamma_i\,\frac{1}{1+i}\right)}{\sum_{i=0}^{\infty}\gamma_i\,\frac{1}{1+i}}$$
$$=-\left[\sum_{i=1}^{\infty}\gamma_i\,\left(1+\frac{2\,i}{n}\right)\right]\int_0^{\bar{y}}\left(\frac{y}{\bar{y}}\right)^i f(y)\,dy$$

or

$$\frac{\gamma_0\left(\sum_{i=1}^{\infty}\gamma_i\,\frac{1}{1+i}\right)-\left(\sum_{i=1}^{\infty}\gamma_i\right)\gamma_0}{\sum_{i=0}^{\infty}\gamma_i\,\frac{1}{1+i}}=-\left[\sum_{i=1}^{\infty}\gamma_i\,\left(1+\frac{2\,i}{n}\right)\right]\int_0^{\bar{y}}\left(\frac{y}{\bar{y}}\right)^i f(y)\,dy$$

and using that $\gamma_0=1$ and rearranging:

$$\sum_{i=1}^{\infty}\frac{\gamma_i\,\frac{1}{1+i}}{\sum_{j=0}^{\infty}\gamma_j\,\frac{1}{1+j}}\,i=\left[\sum_{i=1}^{\infty}\gamma_i\,\left(1+\frac{2\,i}{n}\right)\right]\int_0^{\bar{y}}\left(\frac{y}{\bar{y}}\right)^i f(y)\,dy \tag{62}$$

Using the expression for f, and solving the integrals of terms by term we have:

$$\sum_{j=1}^{\infty}\frac{\gamma_j\,\frac{1}{1+j}}{\sum_{s=0}^{\infty}\gamma_s\,\frac{1}{1+s}}\,j=\sum_{j=1}^{\infty}\gamma_j\,\left(1+\frac{2\,j}{n}\right)\times \tag{63}$$
$$\left(\left[\frac{\left(\frac{\lambda\bar{y}}{2\sigma^2}\right)^{\left(\frac{n}{2}-1\right)}\sum_{i=0}^{\infty}\beta_{i,\frac{n}{2}-1}\,\frac{\bar{y}}{\frac{n}{2}+i+j}\,\left(\frac{\lambda\bar{y}}{2\sigma^2}\right)^i}{\left(\frac{\lambda\bar{y}}{2\sigma^2}\right)^{\left(\frac{n}{2}-1\right)}\sum_{i=0}^{\infty}\beta_{i,\frac{n}{2}-1}\,\left(\frac{\lambda\bar{y}}{2\sigma^2}\right)^i}-\frac{\sum_{i=0}^{\infty}\beta_{i,1-\frac{n}{2}}\,\frac{\bar{y}}{i+1+j}\,\left(\frac{\lambda\bar{y}}{2\sigma^2}\right)^i}{\sum_{i=0}^{\infty}\beta_{i,1-\frac{n}{2}}\,\left(\frac{\lambda\bar{y}}{2\sigma^2}\right)^i}\right]\Bigg/\right.$$
$$\left.\left[\frac{\left(\frac{\lambda\bar{y}}{2\sigma^2}\right)^{\left(\frac{n}{2}-1\right)}\sum_{i=0}^{\infty}\beta_{i,\frac{n}{2}-1}\,\frac{\bar{y}}{\frac{n}{2}+i}\,\left(\frac{\lambda\bar{y}}{2\sigma^2}\right)^i}{\left(\frac{\lambda\bar{y}}{2\sigma^2}\right)^{\left(\frac{n}{2}-1\right)}\sum_{i=0}^{\infty}\beta_{i,\frac{n}{2}-1}\,\left(\frac{\lambda\bar{y}}{2\sigma^2}\right)^i}-\frac{\sum_{i=0}^{\infty}\beta_{i,1-\frac{n}{2}}\,\frac{\bar{y}}{i+1}\,\left(\frac{\lambda\bar{y}}{2\sigma^2}\right)^i}{\sum_{i=0}^{\infty}\beta_{i,1-\frac{n}{2}}\,\left(\frac{\lambda\bar{y}}{2\sigma^2}\right)^i}\right]\right)$$

canceling the values of $\bar{y}$, and defining

$$\xi_i=\frac{1}{i!\,\Gamma\left(i+\frac{n}{2}\right)}\left(\frac{\lambda\bar{y}}{2\sigma^2}\right)^{\left(\frac{n}{2}+i-1\right)}\quad\text{and}\quad\rho_i=\frac{1}{i!\,\Gamma\left(i+2-\frac{n}{2}\right)}\left(\frac{\lambda\bar{y}}{2\sigma^2}\right)^i$$

$$\sum_{j=1}^{\infty} \frac{\gamma_j \, \frac{1}{1+j}}{\sum_{s=0}^{\infty} \gamma_s \, \frac{1}{1+s}} \, j = \sum_{j=1}^{\infty} \gamma_j \, \left(1 + \frac{2\,j}{n}\right) \times \tag{64}$$
$$\left(\left[\frac{\sum_{i=0}^{\infty} \xi_i \, \frac{1}{\frac{n}{2}+i+j}}{\sum_{i=0}^{\infty} \xi_i} - \frac{\sum_{i=0}^{\infty} \rho_i \, \frac{1}{i+1+j}}{\sum_{i=0}^{\infty} \rho_i} \right] \Big/ \left[\frac{\sum_{i=0}^{\infty} \xi_i \, \frac{1}{\frac{n}{2}+i}}{\sum_{i=0}^{\infty} \xi_i} - \frac{\sum_{i=0}^{\infty} \rho_i \, \frac{1}{i+1}}{\sum_{i=0}^{\infty} \rho_i} \right] \right)$$

S Detailed Proof. of Proposition 7.

First we turn to the steady state firm's problem considered in Section 3.2. In that firm's problem we use the same discount rate r for any inflation rate μ. The reason for this is that the period return function is itself normalized by nominal wages which we assume that growth at a constant rate μ and that the nominal rate is equal to $r + \mu$, so that these two effect cancel. The price gap p_i is a real quantity, the difference between the ideal markup and the current markup, and has drift equal to minus the inflation rate due to the increase in the nominal wages. The period return is still $B||p||^2 \equiv B\,y$, but each of the product's price gap evolve as $\mathrm{d}p_i(t) = -\mu\,\mathrm{d}t + \sigma\,\mathrm{d}W_i(t)$. In this problem it is not longer true that y is sufficient to index the state of the firm's problem, since the distribution of $y(t+dt)$ cannot be computed only knowing $y(t)$. While in Alvarez and Lippi (2014) we show that one can take the state to be (y, z) where z is the sum of the price gaps: $z = \sum_{i=1}^{n} p_i$, for the arguments here we keep the entire price gap vector $p \in \mathbb{R}^n$ as the state. In this case the inaction set is no longer a hyper-sphere, nor is the optimal return point to set a zero price gap for each of the products. We let $\mathcal{I}(\mu) \subset \mathbb{R}^n$ be the inaction set –so the firm adjust only if it receives a free adjustment opportunity or if it exist the inaction set. We regard $\mathcal{I}(z)$ as a correspondence parametrized by μ, and let $\hat{p}(\mu) \in \mathbb{R}^n$ be the optimal return point -which is identical across all products- a function parametrized by μ. Note that for any rectangle $\subset\mathbb{R}^n$ the uncontrolled price gaps satisfy that $\Pr\{p(t) - p(0) \in \mathfrak{p} \,|\, \mu\} = \Pr\{-(p(t) - p(0)) \in \mathfrak{p} \,|\, -\mu\}$. This equality uses that the increments of a standard brownian motion are normally distributed. Using this property, and the symmetry around zero of the period return function, it is easy to show that $\hat{p}(\mu) = -\hat{p}(\mu)$. Also, one can see that if $p \in \mathcal{I}(\mu)$ then it must be the case that $-p \in \mathcal{I}(-\mu)$. From these two properties of the decision rules one concludes that $N(\Delta p_i)(\mu)$ and that any even centered moment of the distribution of the price changes, and hence its ratio such as kurtosis $Kur(\Delta p_i; \mu)$, is symmetric around $\mu = 0$. The same property is shown in Alvarez, Lippi, and Paciello (2011) for a closely related model. Likewise, the (negative) symmetry of $\mathcal{M})(\delta, \mu)$ follows by considering first the invariant distribution of price gaps, and then the dynamics of each one. For the invariant distribution of price gaps as defined in Section D, whose density is denoted by $g(p; \mu)$, we note that $g(p; \mu) = g(-p; -\mu)$ –where we now indexed the density only by the inflation rate μ, allowing the optimal decision rule to change with it. Following the same steps we can construct the impulse response of prices $\mathcal{P}(t, \delta; \mu)$ which we index in the same way as the density. We define this impulse response as the change in price level t periods after a once and for all shock δ to the path of the level of money that has occurred to an economy starting at the steady state distribution of price gaps. The price level is in $\mathcal{P}(\delta, t; \mu)$ is measured relative to what the prices would have been absence of a shock, where they would have been rising at a constant rate μ. Using the results previously established we have: $\mathcal{P}(t, -\delta; -\mu) = -\mathcal{P}(t, \delta; \mu)$. Using this property of the impulse response of the price level into definition of $\mathcal{M}$ in equation (12), we obtain the desired (negative) symmetry of this function.

Second, we sketch the differences in the GE set-up when $\mu \neq 0$. In this case the same arguments yields that both nominal interest rates and wages growth at a constant rate μ independently of the distribution of prices at time zero. Additionally, the nominal profit function of the firm, once we replace the first order condition for the households for consumption, labor, and money, can be written as a function of the price gap (i.e. the deviation relative to the markup that maximizes static profits) and the period nominal wages. Hence, one can approximate the real profits (deflated by the money supply) in the same way as with zero inflation, obtaining the same second order approximation. Finally, the result in Proposition 7 in Alvarez and Lippi (2014) which states that GE feedback effects are of order higher than second order in the firm's problem applies almost with no changes.

T Algebraic details for the Proof of Proposition 9

To compute the probabilities $P(t|i)$ notice that

$$P_1\,(t + dt\,|\,i\,) = (1 - \theta_1 dt)\;P_1\,(t\,|\,i\,) + \theta_0\,dt\;[1 - P_1\,(t\,|\,i\,)]$$

for $i \in \{0, 1\}$. Taking a Taylor expansion in $P_1\,(t + dt\,|\,1\,)$, dividing by dt, canceling terms we get:

$$P_1'\,(t\,|\,i\,) = -\,(\theta_1 + \theta_0)\;P_1\,(t\,|\,i\,) + \theta_0$$

The solution of this o.d.e. is:

$$P_1\,(t\,|\,i\,) = \frac{\theta_0}{\theta_0 + \theta_1} + B\,e^{-(\theta_0+\theta_1)t}$$

for some constant B. Evaluating this solution at $t = 0$ we have:

$$P_1\,(0\,|\,i\,) = \frac{\theta_0}{\theta_0 + \theta_1} + B \text{ or } B = P_1\,(0\,|\,i\,) - \frac{\theta_0}{\theta_0 + \theta_1}$$

Thus the solution is

$$P_1\left(t\,|\,i\,\right) = \frac{\theta_0}{\theta_0+\theta_1} + \left[P_1\left(0\,|\,i\,\right) - \frac{\theta_0}{\theta_0+\theta_1}\right]\, e^{-(\theta_0+\theta_1)t}$$

By definition we have

$$P_1\left(0\,|\,0\,\right) = 0 \;\text{ and }\; P_1\left(0\,|\,1\,\right) = 1$$

which gives the desired expressions.

Next , notice that:

$$\mathbb{E}_0\left[\sigma(t)^2|u(0)=i\right] = \sigma_0^2\,\left[1-P_1(t|i)\right] + \sigma_1^2\,P_1(t|i)$$

and

$$\mathbb{E}_0\left[\sigma(t)^2\right] = \mathbb{E}_0\left[\sigma(t)^2|u(0)=1\right]\frac{\theta_0}{\theta_0+\theta_1} + \mathbb{E}_0\left[\sigma(t)^2|u(0)=0\right]\frac{\theta_1}{\theta_0+\theta_1} \quad .$$

Grouping all terms:

$$\mathbb{E}_0\left[\sigma(t)^2\right] = \sigma_0^2 + \left(\sigma_1^2-\sigma_0^2\right)\left[P_1\left(t\,|\,0\right)\frac{\theta_1}{\theta_1+\theta_0} + P_1\left(t\,|\,1\right)\frac{\theta_0}{\theta_1+\theta_0}\right] \quad .$$

We can now use the expressions for $P_1\left(t\,|i\right)$ to obtain:

$$\mathbb{E}_0\left[\sigma(t)^2\right] = \sigma_0^2 + \\ \left(\sigma_1^2-\sigma_0^2\right)\left[\frac{\theta_0}{\theta_0+\theta_1}\left[1-e^{-(\theta_0+\theta_1)t}\right]\frac{\theta_1}{\theta_1+\theta_0} + \frac{\theta_0}{\theta_0+\theta_1}\left[1+\frac{\theta_1}{\theta_0}e^{-(\theta_0+\theta_1)t}\right]\frac{\theta_0}{\theta_1+\theta_0}\right] \quad .$$

Note that we can take common factor $e^{-(\theta_0+\theta_1)t}$ in the right hand side and obtain equation (31).

We also have, using the law of iterated expectations, for $0 \le s \le t \le T$:

$$k(t,s) = \mathbb{E}_0\left[\,\sigma(t)^2\,\sigma(s)^2\,\right] = \mathbb{E}_0\left[\,\sigma(s)^2\,\mathbb{E}\left[\sigma(t)^2\,|\,\sigma^2(s)\right]\right]$$

which we can write as:

$$\mathbb{E}_0\left[\,\sigma(t)^2\,\sigma(s)^2\,\right] = \mathbb{E}_0\left[\,\bar{\sigma}_i^2\,\mathbb{E}\left[\sigma(t)^2\,|\,u(s)=i\right]\right]$$

for which we have the inner expectation:

$$\mathbb{E}\left[\sigma(t)^2\,|\,u(s)=i\right] = \sigma_0^2 + \left(\sigma_1^2-\sigma_0^2\right)P_1\left(t-s\,|\,i\right)$$

and we also have

$$\mathbb{E}_0\left[\,\sigma(t)^2\,\sigma(s)^2\,\right] = \mathbb{E}\left[\,\sigma(t)^2\,\sigma(s)^2\,|u(0)=1\right]\frac{\theta_0}{\theta_0+\theta_1} + \mathbb{E}\left[\,\sigma(t)^2\,\sigma(s)^2\,|u(0)=0\right]\frac{\theta_1}{\theta_0+\theta_1} \quad .$$

Thus we can write (after some algebra):

$$\begin{aligned}
&\mathbb{E}\left[\,\sigma(t)^2\,\sigma(s)^2\,|u(0)=i\right] \\
&= \sigma_0^2\,\left[\sigma_0^2 + \left(\sigma_1^2-\sigma_0^2\right)P_1\left(t-s\,|\,0\right)\right]\left(1-P_1\left(s\,|\,i\right)\right) + \sigma_1^2\,\left[\sigma_0^2 + \left(\sigma_1^2-\sigma_0^2\right)P_1\left(t-s\,|\,1\right)\right]\,P_1\left(s\,|\,i\right) \quad .
\end{aligned}$$

Finally, taking expected values for the initial $u(0)$, and using the formula above, we get:

$$\begin{aligned}
&k(t,s) \\
&= \sigma_0^2\,\left[\sigma_0^2 + \left(\sigma_1^2-\sigma_0^2\right)P_1\left(t-s\,|\,0\right)\right]\left[1-P_1\left(s\,|\,1\right)\frac{\theta_0}{\theta_0+\theta_1} - P_1\left(s\,|\,0\right)\frac{\theta_1}{\theta_0+\theta_1}\right] \\
&+ \sigma_1^2\,\left[\sigma_0^2 + \left(\sigma_1^2-\sigma_0^2\right)P_1\left(t-s\,|\,1\right)\right]\,\left[P_1\left(s\,|\,0\right)\frac{\theta_1}{\theta_0+\theta_1} + P_1\left(s\,|\,1\right)\frac{\theta_0}{\theta_1+\theta_1}\right]
\end{aligned}$$

and using the expressions for the probabilities, or the definition of ergodicity,

$$P_1\left(s\,|\,0\right)\frac{\theta_1}{\theta_0+\theta_1} + P_1\left(s\,|\,1\right)\frac{\theta_0}{\theta_1+\theta_1} = \frac{\theta_0}{\theta_1+\theta_0}$$

we get:

$$k(t,s) = \sigma_0^2\,\left[\sigma_0^2 + \left(\sigma_1^2-\sigma_0^2\right)P_1\left(t-s\,|\,0\right)\right]\frac{\theta_1}{\theta_1+\theta_0} + \sigma_1^2\,\left[\sigma_0^2 + \left(\sigma_1^2-\sigma_0^2\right)P_1\left(t-s\,|\,1\right)\right]\,\frac{\theta_0}{\theta_1+\theta_0} \quad .$$

Replacing the expression for $P_1\left(t-s\,|\,1\right)$ and rearranging we get equation (32).

U A model with two point non-zero random fixed cost

This version of the model assumes that with probability λ per unit of time the menu cost is smaller than the regular adjustment, namely that it costs $b\,\psi$ with $b \in (0,1)$. For simplicity we focus here on a model with one product, i.e. $n=1$. The introduction of a small (but non zero) adjustment cost will imply that this model will have a lower bound for the size of price changes $\underline{p} > 0$, such that no $|\Delta p_i| < \underline{p}$ will be observed. in spite of this important change, which may be important to fit cross section data, we stress that the formula in equation (1) continues to hold.

Firm's problem. The firm's optimal policy now involves two thresholds: $0 < \underline{p} < \bar{p}$. If the price gap is small, i.e. if $|p| \in [0, \underline{p}]$ the firm optimally decides not to adjust the price, even if an opportunity for cheap adjustment occurs. If the price gap is large, i.e. if $|p| \in [\underline{p}, \bar{p})$, the firm adjusts the price only if a cheap adjustment opportunity arises. As in the case where $b = 0$, the firm adjust its price the first time that $|p|$ reaches $\bar{p}$.

Given the values of two thresholds $\underline{p}, \bar{p}$, the value function v can be describe as two functions holding in each segment, as follows:

$$
\begin{aligned}
r\, v_0(p) &= B\, p^2 + \frac{\sigma^2}{2}\, v_0''(p), \quad \text{for } p \in [0, \underline{p}]\ , \\
r\, v_1(p) &= B\, p^2 + \lambda \left[v_0(0) + b\psi - v_1(p) \right] + \frac{\sigma^2}{2}\, v_1''(p), \quad \text{for } p \in [\underline{p}, \bar{p}]
\end{aligned}
$$

where we use that the optimal return point upon adjustment is $v_0(0)$ and where used that by symmetry $v_i(p) = v_i(-p)$ for $i = 0, 1$.

The value function can be expressed as the sum of a particular solution and two solutions multiplied by constants K_0 and K_1 and the two parameters $0 < \underline{p}, \bar{p}$. The value function has the following boundary conditions $v_0(\underline{p}) = v_1(\underline{p})$ and $v_0(0) + \psi = v_1(\bar{p})$, as well as the smooth pasting conditions $v_0'(\underline{p}) = v_1'(\underline{p})$ and $0 = v_1'(\bar{p})$. Using the four boundary conditions one solve for both the value function (i.e. the constants K_i) and the thresholds $\underline{p}, \bar{p}$. We give the details in Appendix U.1 and Appendix U.2.

Frequency of price changes. To find the frequency of price changes we first introduce the expected time to adjustment function $T(p)$. This function obeys the following ODE:

$$
0 = 1 + \frac{\sigma^2}{2} T_0''(p) \quad \text{for } \ 0 < |p| \leq \underline{p} \quad \text{and} \quad \lambda T_1(p) = 1 + \frac{\sigma^2}{2} T_1''(p) \quad \text{for } \underline{p} < |p| \leq \bar{p}
$$

with $T_i(p) = T_i(-p)$, and boundary conditions $T_0(\underline{p}) = T_1(\underline{p})$, $T_0'(\underline{p}) = T_1'(\underline{p})$ and $T_1(\bar{p}) = 0$. Thus

$$
T_0(p) = J - \frac{p^2}{\sigma^2} \ \text{ and } \ T_1(p) = \frac{1}{\lambda} + K e^{\varphi |p|} + L e^{-\varphi |p|}
$$

where the J, K, L are constant to be determined using the boundary conditions, and where $\varphi = \sqrt{2\,\lambda/\sigma^2}$. Thus, given thresholds $\underline{p}, \bar{p}$, solving for the function T boils down to solve three linear equations in three unknowns as detailed in Appendix U.4. In particular the average number of adjustment per period is simply:

$$
N\left(\Delta p_i\right) = \frac{1}{T_0(0)} = \frac{1}{J}\,, \tag{65}
$$

Kurtosis of price changes. To measure the steady state kurtosis of price changes, we first solve for the density function for the price gaps $g(p) \in [0, \bar{p}]$. This density solves

$$\begin{aligned} 0 &= g_0''(p) \text{ for } 0 \leq |p| \leq \underline{p} \quad \text{and} \quad 0 = -\frac{2\lambda}{\sigma^2} g_1(p) + g_1''(p) \text{ for } \underline{p} < |p| \leq \bar{p} \qquad \text{or} \\ g_0(p) &= C_1 + C_2 \, |p| \text{ for } 0 \leq |p| \leq \underline{p} \quad \text{and} \quad g_1(p) = C_3 e^{\varphi |p|} + C_4 e^{-\varphi |p|} \text{ for } \underline{p} \leq |p| \leq \bar{p} \end{aligned}$$

where the 4 constants solve the 4 equations $g_0(\underline{p}) = g_1(\underline{p})$, $g_0'(\underline{p}) = g_1'(\underline{p})$, $g_1(\bar{p}) = 0$ and $1/2 = \int_0^{\underline{p}} g_0(p) \, \mathrm{d}p + \int_{\underline{p}}^{\bar{p}} g_1(p) \, \mathrm{d}p$ which use that the density is differentiable. Given $\underline{p}, \bar{p}$ the solution boils down to solve four linear equations in four unknowns, as detailed in Appendix U.5.

Then using that only the fraction $2 \int_{\underline{p}}^{\bar{p}} g_1(p) \mathrm{d}p$ of cheap adjustment opportunities will trigger an actual price change, the distribution of (non-zero) price changes $p \in [-\bar{p}, -\underline{p}] \cup [\underline{p}, \bar{p}]$ is symmetric and is given by (we only report the formulas for $x > 0$) . Thus the distribution of (positive) price changes is

$$\text{Price changes} \sim \begin{cases} \text{density for a price change of size } p \in [\underline{p}, \bar{p}) & : \frac{\lambda}{N_a} g_1(p) \\ \text{mass point at } \bar{p} & : \frac{1}{2} - \frac{\lambda}{N_a} \int_{\underline{p}}^{\bar{p}} g_1(p) \mathrm{d}p \end{cases}$$

The $j - th$ moment of price changes for j even is

$$\mathbb{E}\left(\Delta p^j\right) = \frac{\lambda}{N_a} 2 \int_{\underline{p}}^{\bar{p}} x^j \, g_1(p) \, \mathrm{d}p + \left(1 - \frac{\lambda \, 2 \int_{\underline{p}}^{\bar{p}} g_1(p) \mathrm{d}p}{N_a} \right) \bar{p}^j$$

Using that $Var(\Delta p) \, N(\Delta p) = \sigma^2$, the kurtosis of price changes is given by:

$$Kur(\Delta p) = \frac{\mathbb{E}(\Delta p^4)}{\left(\sigma^2 / N(\Delta p)\right)^2} . \tag{66}$$

Area under impulse response. To find an expression for $\mathcal{M}'(0)$ we first define the contribution to the area under impulse response of a firm that starts with price gap p. Letting $m(p)$ the integral of the (minus) expected price gap until the first time the firms adjusts its price, and starting the economy with a distribution of price gaps with density f we have

$$\mathcal{M}(\delta) = \int_{-\bar{p}}^{\bar{p}} m(p - \delta) \, g(p) \, dp \tag{67}$$

and differentiating it:

$$\mathcal{M}'(0) = -\int_{-\bar{p}}^{\bar{p}} m'(p) \, g(p) \, dp \tag{68}$$

To obtain the solution for m we consider two functions in each segments which solves:

$$0 = -p + \frac{\sigma^2}{2} m_0''(p) \text{ for } 0 \le p \le \underline{p} \tag{69}$$

$$\lambda\, m_1(p) = -p + \frac{\sigma^2}{2} m_1''(p) \text{ for } \underline{p} \le p \le \bar{p} \tag{70}$$

The boundary conditions are that these functions meet in a continuously differentiable manner in the lower boundary, i.e. $m_0(\underline{p}) = m_1(\underline{p})$, $m_0'(\underline{p}) = m_1'(\underline{p})$, and that a price change occurs at the upper boundary, i.e. $m_1(\bar{p}) = 0$. The solution, with three constant of integration is:

$$m_0(p) = A_1\, p + \frac{p^3}{3\,\sigma^2} \tag{71}$$

$$m_1(p) = -\frac{p}{\lambda} + A_2\, e^{p\varphi} + A_3\, e^{-p\varphi} \tag{72}$$

Thus, given $\underline{p}, \bar{p}$ boils down to solving three linear equations in three unknowns, as detailed in Appendix U.6.

Hence, given any pair $(\underline{p}, \bar{p})$ we can find the solution for the density g, the solution to the function m and compute:

$$\mathcal{M}'(0) = -2 \int_0^{\bar{p}} m'(p)\; g(p)\, dp$$

Likewise, given any pair $(\underline{p}, \bar{p})$, we can find the solution for g, $N\,(\Delta p)$ and compute $Kur\,(\Delta p)$ as in equation (66). In Appendix U.7 we collect the solutions as function of the thresholds $(\underline{p}, \bar{p})$ and constants $(A_1, A_2, A_3, J, C_3, C_4)$. From this one can easily compute both expressions and check the equality in

$$\mathcal{M}'(0) = \frac{Kur\,(\Delta p)}{6\,N\,(\Delta p)}\ .$$

U.1 Solution of ode for value function in inaction

$$v_0(p) = \frac{B\,p^2}{r} + \frac{B\,\sigma^2}{r^2} + K_0 \left(e^{p\sqrt{\frac{2r}{\sigma^2}}} + e^{-p\sqrt{\frac{2r}{\sigma^2}}} \right)$$

$$v_1(p) = \frac{B\,p^2 + \lambda\,(v_0(0) + b\psi)}{\lambda + r} + \frac{B\,\sigma^2}{(\lambda + r)^2} + K_1 \left(e^{p\sqrt{\frac{2(\lambda+r)}{\sigma^2}}} + e^{-p\sqrt{\frac{2(\lambda+r)}{\sigma^2}}} \right)$$

U.2 Solution for value function

Note that smooth pasting $v_1'(\bar{p}) = 0$ gives

$$0 \quad = \quad \frac{2B\bar{p}}{\lambda + r} + K_1 \sqrt{\frac{2(\lambda + r)}{\sigma^2}} \left(e^{\bar{p}\sqrt{\frac{2(\lambda+r)}{\sigma^2}}} - e^{-\bar{p}\sqrt{\frac{2(\lambda+r)}{\sigma^2}}} \right)$$

or K_1 as function of $\bar{p}$

$$K_1 \quad = \quad \frac{2B\bar{p}}{\lambda + r} \left[\sqrt{\frac{2(\lambda + r)}{\sigma^2}} \left(e^{-\bar{p}\sqrt{\frac{2(\lambda+r)}{\sigma^2}}} - e^{\bar{p}\sqrt{\frac{2(\lambda+r)}{\sigma^2}}} \right) \right]^{-1} \tag{73}$$

Using $v_0(0) = \frac{B\sigma^2}{r^2} + 2K_0$ and value matching $v_0(0) + \psi = v_1(\bar{p})$ gives

$$\frac{r}{\lambda + r} v_0(0) + \psi \quad = \quad \frac{\lambda b \psi + B\bar{p}^2}{\lambda + r} + \frac{B\sigma^2}{(\lambda + r)^2} + K_1 \left(e^{\bar{p}\sqrt{\frac{2(\lambda+r)}{\sigma^2}}} + e^{-\bar{p}\sqrt{\frac{2(\lambda+r)}{\sigma^2}}} \right)$$

or K_0 as function of $\bar{p}$

$$2rK_0 \quad = \quad B\bar{p}^2 - (\lambda(1-b) + r)\psi - \frac{\lambda B\sigma^2}{r(\lambda + r)} + (\lambda + r)K_1 \left(e^{\bar{p}\sqrt{\frac{2(\lambda+r)}{\sigma^2}}} + e^{-\bar{p}\sqrt{\frac{2(\lambda+r)}{\sigma^2}}} \right) \tag{74}$$

Value matching at $\underline{p}$ gives

$$\frac{B\underline{p}^2}{r} + \frac{B\sigma^2}{r^2} + K_0 \left(e^{\underline{p}\sqrt{\frac{2r}{\sigma^2}}} + e^{-\underline{p}\sqrt{\frac{2r}{\sigma^2}}} \right) \quad = \quad \frac{B\underline{p}^2 + \lambda\left(v_0(0) + b\psi\right)}{\lambda + r} + \frac{B\sigma^2}{(\lambda + r)^2} + K_1 \left(e^{\underline{p}\sqrt{\frac{2(\lambda+r)}{\sigma^2}}} + e^{-\underline{p}\sqrt{\frac{2(\lambda+r)}{\sigma^2}}} \right)$$

or an equation implicitly defining $\underline{p}$ in terms of $\bar{p}$

$$\frac{B\underline{p}^2\lambda}{r(r + \lambda)} + \frac{B\sigma^2\lambda}{(\lambda + r)^2 r} + K_0 \left(e^{\underline{p}\sqrt{\frac{2r}{\sigma^2}}} + e^{-\underline{p}\sqrt{\frac{2r}{\sigma^2}}} - \frac{2\lambda}{\lambda + r} \right) \quad = \quad \frac{\lambda b \psi}{\lambda + r} + K_1 \left(e^{\underline{p}\sqrt{\frac{2(\lambda+r)}{\sigma^2}}} + e^{-\underline{p}\sqrt{\frac{2(\lambda+r)}{\sigma^2}}} \right)$$

Given these 3 equations implicitly defining $K_0, K_1, \underline{p}$ as function of $\bar{p}$, the smooth pasting at $\underline{p}$ gives one equation in one unknown to solve for $\bar{p}$, namely

$$\left(\frac{2B}{r} - \frac{2B}{r + \lambda} \right) \underline{p} + \sqrt{\frac{2r}{\sigma^2}} K_0 \left(e^{\underline{p}\sqrt{\frac{2r}{\sigma^2}}} - e^{-\underline{p}\sqrt{\frac{2r}{\sigma^2}}} \right) = \sqrt{\frac{2(\lambda + r)}{\sigma^2}} K_1 \left(e^{\underline{p}\sqrt{\frac{2(\lambda+r)}{\sigma^2}}} - e^{-\underline{p}\sqrt{\frac{2(\lambda+r)}{\sigma^2}}} \right)$$

U.3 Value function approximation

Recall

$$\begin{aligned}
v_0(p) &= \frac{Bp^2}{r} + \frac{B\sigma^2}{r^2} + K_0\left(e^{p\sqrt{\frac{2r}{\sigma^2}}} + e^{-p\sqrt{\frac{2r}{\sigma^2}}}\right) \\
v_1(p) &= \frac{Bp^2 + \lambda\left(v_0(0) + b\psi\right)}{\lambda + r} + \frac{B\sigma^2}{(\lambda + r)^2} + K_1\left(e^{p\sqrt{\frac{2(\lambda+r)}{\sigma^2}}} + e^{-p\sqrt{\frac{2(\lambda+r)}{\sigma^2}}}\right)
\end{aligned}$$

We approximate the value functions $v_0(p), v_1(p)$ using a fourth order expansion around $p = 0$. We get

$$\begin{aligned}
v_0(p) &= \frac{B\sigma^2}{r^2} + 2K_0 + \left(\frac{B}{r} + K_0\varphi_0^2\right)\, p^2 + \frac{K_0}{12}\varphi_0^4\, p^4 \\
v_1(p) &= \frac{\lambda(v_0(0) + b\psi)}{\lambda + r} + \frac{B\sigma^2}{(\lambda + r)^2} + 2K_1 + \left(\frac{B}{\lambda + r} + K_1\varphi_1^2\right)\, p^2 + \frac{K_1}{12}\varphi_1^4\, p^4 \\
&\text{where } \varphi_0 \equiv \sqrt{\frac{2r}{\sigma^2}} \quad \text{and} \quad \varphi_1 \equiv \sqrt{\frac{2(\lambda + r)}{\sigma^2}}
\end{aligned}$$

The smooth pasting at $\underline{p}$, namely $v_0'(\underline{p}) - v_1'(\underline{p}) = 0$, gives

$$p\left[\left(\frac{B}{r} + K_0\varphi_0^2\right) - \left(\frac{B}{\lambda + r} + K_1\varphi_1^2\right) + \left(K_0\varphi_0^4 - K_1\varphi_1^4\right)\frac{p^2}{6}\right] = 0$$

which gives

$$\underline{p} = \pm\sqrt{\frac{\left(\frac{B}{\lambda+r} + K_1\varphi_1^2\right) - \left(\frac{B}{r} + K_0\varphi_0^2\right)}{\left(K_0\varphi_0^4 - K_1\varphi_1^4\right)/6}}$$

Similarly smooth pasting at $\bar{p}$ gives

$$\bar{p} = \pm\sqrt{\frac{\left(\frac{B}{\lambda+r} + K_1\varphi_1^2\right)}{-K_1\varphi_1^4/6}}$$

U.4 Boundary conditions for T_i

We have the following three linear equations for T_i:

$$\begin{aligned} -\frac{1}{\lambda} &= Ke^{\varphi\bar{p}} + Le^{-\varphi\bar{p}} \\ -\frac{2\,\underline{p}}{\sigma^2} &= \varphi\left(Ke^{\varphi\underline{p}} - Le^{-\varphi\underline{p}}\right) \\ J &= \frac{\left(\underline{p}\right)^2}{\sigma^2} + \frac{1}{\lambda} + Ke^{\varphi\underline{p}} + Le^{-\varphi\underline{p}} \end{aligned}$$

U.5 Density function

The 4 unknowns of the density function, using $g_1(\bar{p}) = 0$ and $g_0'(\underline{p}) = g_1'(\underline{p})$, give

$$C_3 = -C_4 e^{-2\varphi\bar{p}} \quad \text{and} \quad C_2 = -C_4\varphi\left(e^{-2\varphi\bar{p}+\varphi\underline{p}} + e^{-\varphi\underline{p}}\right)$$

Next, using $g_0(\underline{p}) = g_1(\underline{p})$ gives

$$C_1 = -C_2\underline{p} - C_4\left(e^{-2\varphi\bar{p}+\varphi\underline{p}} - e^{-\varphi\underline{p}}\right) = C_4\left[e^{-2\varphi\bar{p}+\varphi\underline{p}}\left(\varphi\underline{p} - 1\right) + e^{-\varphi\underline{p}}\left(\varphi\underline{p} + 1\right)\right]$$

Finally we solve for C_4 by imposing $1/2 = \int_0^{\underline{p}} g_0(p)\,\mathrm{d}p + \int_{\underline{p}}^{\bar{p}} g_1(p)\,\mathrm{d}p$ i.e.

$$\frac{1}{2} = C_1\underline{p} + \frac{1}{2}C_2\underline{p}^2 + \frac{1}{\varphi}\left[C_3\left(e^{\varphi\bar{p}} - e^{\varphi\underline{p}}\right) - C_4\left(e^{-\varphi\bar{p}} - e^{-\varphi\underline{p}}\right)\right]$$

or, substituting the expressions,

$$\begin{aligned} \frac{1}{2C_4} &= \left[e^{-2\varphi\bar{p}+\varphi\underline{p}}\left(\varphi\underline{p} - 1\right) + e^{-\varphi\underline{p}}\left(\varphi\underline{p} + 1\right)\right]\underline{p} - \frac{1}{2}\varphi\left(e^{-2\varphi\bar{p}+\varphi\underline{p}} + e^{-\varphi\underline{p}}\right)\underline{p}^2 \\ &\quad -\frac{1}{\varphi}\left[e^{-2\varphi\bar{p}}\left(e^{\varphi\bar{p}} - e^{\varphi\underline{p}}\right) + e^{-\varphi\bar{p}} - e^{-\varphi\underline{p}}\right] \end{aligned}$$

U.6 Equation for the solution of m

The boundary conditions are: $m_1(\bar{p}) = 0$, $m_1(\underline{p}) = m_0(\underline{p})$ and $m_1'(\underline{p}) = m_0'(\underline{p})$. They give a linear system of equations on A_1, A_2, A_3 :

$$0 = -\frac{\bar{p}}{\lambda} + A_2\,e^{\bar{p}\varphi} + A_3\,e^{-\bar{p}\varphi} \tag{75}$$

$$A_1 + \frac{(\underline{p})^2}{\sigma^2} = -\frac{1}{\lambda} + \varphi\,A_2\,e^{\underline{p}\varphi} - \varphi\,A_3\,e^{-\underline{p}\varphi} \tag{76}$$

$$A_1\,\underline{p} + \frac{(\underline{p})^3}{3\,\sigma^2} = -\frac{\underline{p}}{\lambda} + A_2\,e^{\underline{p}\varphi} + A_3\,e^{-\underline{p}\varphi} \tag{77}$$

U.7 Algebraic details for main proposition

For the area under the IRF of output we get:

$$\begin{aligned}\mathcal{M}'(0) = & - \quad 2\int_0^{\underline{p}} \left[A_1 + A_3 \frac{p^2}{\sigma^2}\right] [C_1 + C_2 p] \, \mathrm{d}p \\ & - \quad 2\int_{\underline{p}}^{\bar{p}} \left[-\frac{1}{\lambda} + \varphi \, A_2 \, e^{p\varphi} + \varphi \, A_3 \, e^{-p\varphi}\right] \left[C_3 e^{\varphi p} + C_4 e^{-\varphi p}\right] \mathrm{d}p\end{aligned}$$

For the kurtosis of steady state price changes we get:

$$\begin{aligned}\frac{Kur\,(\Delta p)}{6\,N\,(\Delta p)} &= N\,(\Delta p) \, \frac{\mathbb{E}\,(\Delta p^4)}{6\,\sigma^4} \\ &= \frac{\lambda\,J\,2\,\int_{\underline{p}}^{\bar{p}} p^4 \, \left[C_3 e^{\varphi p} + C_4 e^{-\varphi p}\right] \mathrm{d}p + \left(1 - \frac{\lambda\,2\,\int_{\underline{p}}^{\bar{p}} \left[C_3 e^{\varphi p} + C_4 e^{-\varphi p}\right] \mathrm{d}p}{N_a}\right) \bar{p}^4}{6\,J\,\sigma^4} \\ &= \frac{\lambda\,2}{6\,\sigma^4} \int_{\underline{p}}^{\bar{p}} p^4 \, \left[C_3 e^{\varphi p} + C_4 e^{-\varphi p}\right] \mathrm{d}p + \frac{1}{6\,J\,\sigma^4} - \frac{\lambda\,2}{6\,\sigma^4} \int_{\underline{p}}^{\bar{p}} \bar{p}^4 \, \left[C_3 e^{\varphi p} + C_4 e^{-\varphi p}\right] \mathrm{d}p \\ &= \frac{\lambda\,2}{6\,\sigma^4} \int_{\underline{p}}^{\bar{p}} \left(p^4 - \bar{p}^4\right) \, \left[C_3 e^{\varphi p} + C_4 e^{-\varphi p}\right] \mathrm{d}p + \frac{1}{6\,J\,\sigma^4}\end{aligned}$$