Are State and Time dependent models really different? *

Fernando Alvarez
University of Chicago and NBER

Francesco Lippi
Einaudi Institute for Economics and Finance, University of Sassari, and CEPR

Juan Passadore
Einaudi Institute for Economics and Finance

April 13, 2016

FIRST DRAFT

Abstract

Yes, but only for large monetary shocks. In particular, we show that for a large class of models where shocks have continuous paths, the propagation of a monetary impulse is independent of the nature of the sticky price friction when shocks are *small*. The propagation of large shocks instead depends on the nature of the friction: the impulse response of inflation to monetary shocks is non-linear in state-dependent models, while it is independent of the shock size in time-dependent models. We use data on exchange rate devaluations and inflation for a panel of countries over 1974-2014 to test for the presence of state dependent decision rules. We find evidence of a non-linear effect of exchange rate changes on prices in the sample of flexible-exchange rate countries with low inflation. In particular, we find that large exchange rate changes have larger short term pass through, as implied by state dependent models.

JEL Classification Numbers: E5

Key Words: observation costs, menu cost, inattentiveness, monetary shocks, impulse responses, time dependent, state dependent, exchange rate pass-through.

*First draft, February 2016. Prepared for the NBER Macro Annual meeting, 2016. Part of the research for this paper was sponsored by the ERC advanced grant 324008. Joaquin Saldain and Jean Flemming provided excellent research assistance.

1 Introduction

During the last decade the analysis of new micro data has contributed to advancing the sticky price literature, by challenging existing models and fostering the development of new ones. Indeed, current frontier models are consistent with several cross sectional facts about the size-distribution as well as the timing of price changes uncovered by the micro data. An open issue in this research agenda concerns the nature of, or the appropriate underlying friction used to model, sticky prices. Two alternative assumptions to generate infrequent adjustment of prices involve either a fixed cost, as in Golosov and Lucas (2007) menu-cost model, or a limited information-gathering and information-processing ability, as in Reis (2006) "rational inattentiveness" setup. Following the description used in the literature, we refer to these types of models as "state-dependent" models or "time-dependent" models.[1] Some scholars argue that information frictions will generate stronger real effects of monetary policy shocks, (e.g. Mankiw and Reis (2002); Klenow and Kryvtsov (2008)), but an analytic comparison of the consequences of each of these mechanisms for the transmission of monetary policy shocks has not been developed. Under what circumstances does the nature of the underlying friction matter for the propagation of monetary shocks? What kind of empirical evidence can be used to identify the nature of the underlying friction? This paper casts light on these questions by presenting new theoretical results and some evidence that bears upon such theories.

The first part of the paper formalizes the definition of time-dependent and state-dependent models (TD and SD, respectively), and analyzes the propagation of monetary shocks under the different frictions. The class of models we consider embeds (approximately) several classic sticky price models, such as Taylor (1980); Calvo (1983); Reis (2006); Golosov and Lucas (2007); Nakamura and Steinsson (2010); Midrigan (2011); Bonomo, Carvalho, and Garcia (2010); Bhattarai and Schoenle (2014); Carvalho and Schwartzman (2015) and several other cases that are novel in the literature. All the models considered are characterized by the presence of idiosyncratic shocks with continuous paths. We concentrate on three types of results, which taken together show that what distinguishes state and time dependent models is their reaction to a *large* aggregate shock. The first result, which holds in a very general class of models, characterizes the impact effect on the aggregate price of a common unanticipated permanent shock to the nominal cost of all firms. The second result, which is analyzed for a smaller set of economies but which deals with the general equilibrium effects, characterizes the output response to an unanticipated permanent increase of the money supply for an economy in steady state. We show that for small shocks the nature of the friction is *irrelevant*, i.e.

[1] In the first class of models the firm's decision to adjust prices depends on the state, while in the latter class it depends only on the time elapsed since the last price change.

the results are the same for state and time dependent models.[2] More specifically, in our first result we follow the lead of Caballero and Engel (2007) and analyze the *impact* effect of a monetary shock on inflation, a statistic they refer to as the "Flexibility index". This statistic corresponds to the impact effect (i.e. the initial point) of the impulse response function: the inflation reaction at the time of the shock. We show that the flexibility index is always zero in TD models. More surprisingly, we show that the shock does not have a first order effect on the aggregate price even in SD models, so that for small shocks the impact is approximately zero provided that firms follow an Ss decision rule (possibly multidimensional) and that the shocks faced by the firms follow a diffusion. Our second result extends this irrelevance beyond the impact-effect, by considering the total cumulated *output* response triggered by a small monetary shock (measured by the area under the output impulse response function). For economies with low inflation we show that the total cumulated output response is the same in TD and SD models, provided the models are fit to the same steady state moments. These results are quite robust; we show that they also apply in the presence of moderate rates of steady state inflation.

The third theoretical result highlights a key difference between SD and TD models, which appears when the aggregate shock is large. In TD models the impulse response function of prices at any given horizon is proportional to the size of the shock. Furthermore, as implied by our previous results, the impact effect of the shock on aggregate prices is zero for any shock size. These features imply that the shape of the impulse response does not depend on the size of the shock. Instead, the inherent non-linear nature of decision rules of SD models implies that for aggregate shocks above a minimum size, the economy displays full price flexibility. Thus for SD models the impact effect of the shock depends on their *size*. This prediction suggests a simple test for the nature of the friction behind sticky prices: TD models predict a proportional response in terms of the size of the shock, while SD models predict a non-linear response with respect to the size of the shocks.

The final part of the paper presents an empirical investigation of the hypothesis, inspired by the above theoretical results, that the response of inflation to a monetary shock, particularly on impact, depends on the size of the shock. We follow the international economics literature that studies the pass-through of exchange rate shocks on to prices, as in e.g. Burstein and Gopinath (2014). In particular, we use a panel of monthly data from a large number of countries about CPI inflation and the nominal exchange rate (bilateral exchange rate vs the dollar). To keep in line with the theory we restrict attention to countries

[2] Small shocks are indeed central to several previous analysis both because they naturally emerge as the residuals in a regression, as typical in the empirical VAR literature (e.g. Christiano, Eichenbaum, and Evans (1999, 2005)), or because they provide convenient conditions for analytical approximations (e.g. Caballero and Engel (2007); Alvarez, Le Bihan, and Lippi (2016)).

with moderate inflation in the post Bretton-Woods period. Our preliminary empirical results uncover some evidence of a non-linear pass-through of devaluation on inflation for a sample that excludes countries in a fixed exchange rate regime as classified by Levy-Yeyati and Sturzenegger (2003) and Ilzetzki, Reinhart, and Rogoff (2008). This evidence shows that the inflation response to an exchange rate shock depends on the size of the shock. For instance, in the month following the shock the elasticity of inflation with respect to a 5.5% shock is almost two times larger than the elasticity to a 1% shock. Interestingly, those differences can be noticed only in the first months after the impact, and eventually disappear, consistent with the view that large shocks trigger a faster response of the economy. A similar pattern is found for all countries in the post 1990 sample, and appears to hold with different functional form specifications (piecewise linear, quadratic and cubic, or distributed lags), different controls (e.g. for fixed vs flex exchange rate regime, GDP growth rates). However, we notice that the evidence of the non-linear effect is more noisy when Fixed exchange rate countries are added to the sample.

Novelty and relation to literature. Our analysis is inspired by the work of Klenow and Kryvtsov (2008). Like them we also aim at investigating the nature of the frictions that underlie sticky prices. The two papers however have a different focus. Their pioneering paper mostly focuses on the documentation the micro facts and in assessing the success of several classic models (encompassed by our framework) in matching the cross-sectional data. Their analysis does not investigate how the different models behave in response to an aggregate shock, which instead is the focus of our analysis. We aim to identify the implications of the different frictions for the propagation of aggregate shocks, and provide original analytic results that are useful for a systematic comparison of the two approaches.

Our first result extends the theoretical analysis of Caballero and Engel (2007) about the aggregate flexibility of an economy, which focuses on the impact effect of a monetary shock. We analytically show that in several TD and SD models of the last generation, featuring idiosyncratic shocks, a small monetary shock does not have a first order effect on the aggregate price level. Second, this paper unifies recent results that focus on the full profile of the impulse response function, not just on the impact effect. These results build on, and extend, previous contributions in Alvarez and Lippi (2014); Alvarez, Le Bihan, and Lippi (2016); Alvarez, Lippi, and Paciello (2016). We generalize the previous results by showing that they also hold in settings that feature both state and time dependent frictions as considered by Abel, Eberly, and Panageas (2007, 2013); Alvarez, Lippi, and Paciello (2011); Bonomo, Carvalho, and Garcia (2010) and Bonomo et al. (2016).

While the class of models we analyze is large, we comment next on some models that do not belong to it. For instance in models with no idiosyncratic shocks such as Sheshinski and Weiss (1983), and the classic analysis of monetary shocks in this environment by Caplin and Spulber (1987) and Caplin and Leahy (1991), small monetary shocks have a first order impact effect on inflation. We discuss the relation with these results in detail in Section 4.1. Moreover, our setup with idiosyncratic shocks with continuous paths rules out models where firms are hit by infrequent and large idiosyncratic shocks, as considered by Gertler and Leahy (2008) or Midrigan (2011). While not all of our theoretical results hold as stated for these models, the main idea still applies. In particular, small monetary shocks have no impact effect (i.e. they are second order), but large monetary shocks have a first order effect due to the state dependence of the decision rules.

Our paper also provides some novel empirical analysis of the non-linear passthrough prediction. Even though most of the pass through literature focuses on the magnitude of linear terms (e.g. Burstein and Gopinath (2014), Campa and Goldberg (2005) and Martins (2005)), there at least two papers that test for non-linearities with a specification that is similar to ours. Pollard and Coughlin (2004) study the non-linear response to exchange rate shocks of US industries using Import Prices and find that firms in over half the industries respond asymmetrically. Bussiere (2013) studies the non-linear response to exchange rate shocks using import and export prices of G7 countries and finds evidence, with a specification similar to ours, of a nonlinear response in country by country regressions as well as in panel regressions.

More broadly, we see our paper as a contribution to the burgeoning literature on non-linear effects in macroeconomics. Examples of this literature are the macro-finance models such as Brunnermeier and Sannikov (2014), as well as the models featuring the zero lower bound, such as Fernandez-Villaverde et al. (2015). In these models shocks in different regions of the state space have differential effects, which turns out to be important for policy. Our contribution focuses on the differential effect of shocks according to their size, as in the seminal empirical analysis of fiscal policy by Giavazzi, Jappelli, and Pagano (2000), and the more recent quantitative models as in e.g. Kaplan and Violante (2014). The class of SD models we analyze, like these models, features a size-asymmetry in the economy's response to small and large shocks. The application and the models are, of course, different: we focus on how prices respond to monetary shocks they focus on the consumption response to fiscal shocks. Also, we offer some preliminary evidence of the non-linear pass-through of nominal exchange rate changes on inflation using a large panel of countries.

Organization of the paper. The next section gives a broad non-technical overview of the modeling setup and a summary of the main results. Section 3 describes the setups we use to analyze state-dependent models, time-dependent models, as well as a setting featuring both time and state dependent features. Section 4 outlines the main theoretical results we derive for these economies concerning the propagation of monetary shocks. We first discuss the result on the equivalence between these models in the presence of small monetary shocks. Next we discuss the differences between these models that appear with large shocks. Section 5 presents the empirical analysis.

2 Overview of main results

We start by defining the elements for state-dependent (SD) and time-dependent (TD) setups. A *state dependent* setup is one where price changes occur if the state, given by current profits or markups, attains a critical level. The SD models are characterized by decision rules that depend on the value of the state. In the presence of adjustment costs, necessary to model sticky prices, the state space of the problem is split in two regions, one where inaction is optimal, and another where the firms find it optimal to adjust prices to return the state to a point located well inside this set. Price changes occur when the state reaches the boundary of the inaction region. A *time dependent* model is one where the times between consecutive price changes are statistically independent of the state (e.g. the current markup or profits of the firm). Instead, the time elapsed since the last price change (and potentially the duration of the previous price spell) completely determines the hazard rate of price changes. We show how the decision rules that correspond to TD or SD can be derived from an explicit profit maximization problem in the presence of fixed cost to observing the state ($\psi_o > 0$) or fixed cost to adjusting the nominal price ($\psi_m > 0$), respectively.

Our analysis focuses on the propagation of a permanent unexpected shock δ, measuring the change (in log points) of the nominal marginal cost of all firms, starting from the steady state of an economy with an inflation rate π and idiosyncratic shocks with variance σ^2. There are three theoretical results that we discuss. The first result concerns the impact effect of δ on the price level. For this result we don't need to specify the whole economy, instead we just take a continuum of firms that solve the type of problem described above and that face a common (once and for all) nominal cost shock.[3] We show that for both TD and SD models a small monetary shock δ has a second order effect on the price level $\mathcal{P}(\delta, t)$ on impact, for any $\pi > 0$ provided that $\sigma > 0$. Formally, let $\mathcal{P}(\delta, t)$ be the price level $t \geq 0$ periods after an

[3]The approach is standard and has been used in e.g. Caballero and Engel (2007).

unexpected increase of the money supply of size δ. This implies

$$\mathcal{P}(\delta, t) \;\; = \;\; \Theta(\delta) + \int_0^t \theta(\delta, s) ds \qquad (1)$$

where $\Theta(\delta)$ denotes the impact response of the price level at the time of the shock. In particular we show that in all TD, SD, and mixed models, we have that $\Theta'(0) = 0$, i.e. that there is no first-order effect of the monetary shock on the price level. An illustration of this result can be seen in Figure 1, which plots the response of output to a permanent monetary shock for 3 economies characterized by the same frequency of price changes per year (normalized to unity) and different kurtosis of the size distribution of price changes. It appears that as the 1% shock hits the economy output increases by approximately 1% in all economies, since the CPI does not respond on impact.[4] This result is of interest because it clarifies previous analyses of the impact effect. For instance Caballero and Engel (2007) propose a theoretical characterization of the impact effect, $\Theta'(0)$, which they refer to as the *flexibility index*, as a way to characterize different sticky price models. Thus in a large class of models, with $\sigma > 0$ and $\pi > 0$, the flexibility index is zero.

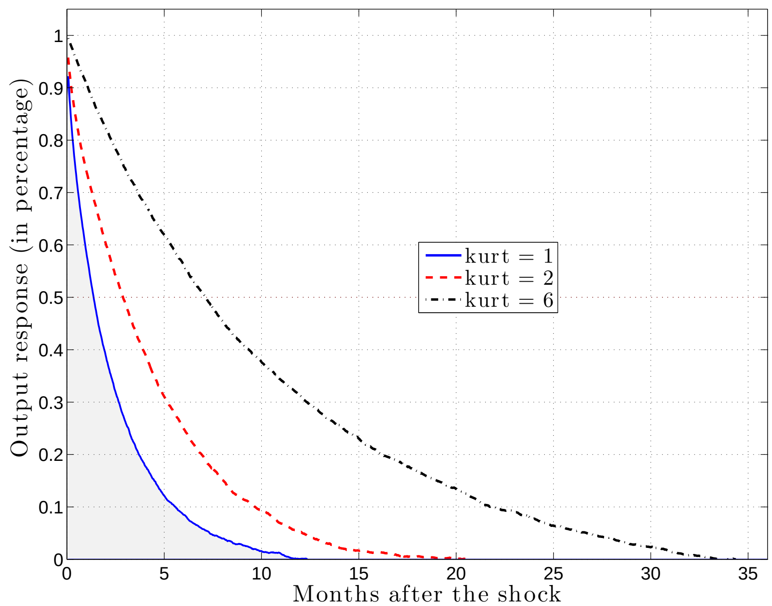
While monetary shocks do not have a first order impact on the aggregate price level in neither TD nor SD models, the reason behind this result is different. For TD models, the distribution of the number of firms adjusting at different times is independent of the aggregate shock. Thus, the aggregate price level does not jump on impact, i.e. $\Theta(\delta) = 0$ for all δ. For SD models the result is due to the fact that there is no "mass" of firms close to the adjustment boundaries (literally, a zero density), which in turn is explained because the boundaries are exit points where all firms adjust.[5] Thus, in SD models $\Theta(\delta)$ is of order δ^2, so small shocks trigger extremely small jumps.

An important property of the impact effect concerns how it changes as a function of inflation, π, relative to the volatility of the idiosyncratic shocks σ. While for $\sigma > 0$ the impact effect Θ is of order δ^2, we notice that the impact effect is increasing with π and that the effect becomes first order as π/σ diverges. The menu cost models of Sheshinski and Weiss (1983) and Caplin and Spulber (1987) illustrate this point: in both models the impact effect $\Theta(\delta)$ is of order δ, the reason is that in these models $\sigma = 0$ and $\pi > 0$, so that the ratio diverges. Thus, since the impact effect is second order but it is increasing in π, in the empirical analysis we will focus on low-inflation countries where the lack of response to small shocks should be easier to detect.

[4] This example assumes a unit elasticity of output to real wages.

[5] Technically this last result depends on the continuous-time and continuous-path nature of the shocks, but its qualitative implications also apply to discrete-time discrete-state versions of this model.

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



The figure represents an economy with $\epsilon = 1$, $N(\Delta p_i) = 1.0$ and $std(\Delta p_i) = 0.10$. The three curves correspond to economies with a steady state kurtosis of the size of price changes equal to 1, 2 and 6, respectively.

The second result goes beyond the analysis of the impact effect and considers a summary measure for the whole profile of the impulse response function. We derive this second result focusing on economies where the steady state inflation equals (or is close to) zero.[6] Moreover, for this result we completely specify a General Equilibrium effect, so the shock is interpreted as a monetary shock. Specifically, the summary statistic that we choose is the area under the output impulse response function following an increase of the money supply of size δ. We denote this magnitude by $\mathcal{M}(\delta)$, e.g. the gray shaded area that appears for illustrative purposes in Figure 1 for three models with kurtosis equal to 1, 2 and 6, respectively. Formally, the cumulative output $\mathcal{M}$ after a shock δ is:

$$\mathcal{M}(\delta) = \frac{1}{\epsilon} \int_0^\infty \left(\delta - \mathcal{P}(\delta, t)\right) \, dt \tag{2}$$

where $\mathcal{P}(\delta, t)$ is the aggregate price level t periods after the shock δ. The argument of the integral gives the aggregate real wages at time t, which are then mapped into output by $1/\epsilon$, a parameter related to the elasticity of the labor supply. Integrating over time gives the total cumulative real output. We find the $\mathcal{M}$ 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.[7] Second for small monetary shocks (like the ones typically considered in the literature) this statistic is completely encoded by a simple formula that involves the frequency of price changes $N(\Delta p_i)$ and the kurtosis of price changes $Kur(\Delta p_i)$.

We show that in state-dependent (SD) and time-dependent (TD) models, as well as in models where both TD and SD features, 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 $1/\epsilon - 1$ we show that for a small monetary shock δ the cumulative output $\mathcal{M}$ is accurately approximated by the following expression

$$\mathcal{M}(\delta) \approx \frac{\delta}{6\,\epsilon} \; \frac{Kur(\Delta p_i)}{N(\Delta p_i)} \quad . \tag{3}$$

An immediate implication of this result is that for small monetary shocks the underlying

[6] Theoretically the result extends to small inflation since in the presence of idiosyncratic shocks the drift has a second order impact on decision rules, such as the frequency of price adjustments (see Alvarez, Lippi, and Paciello (2011) for a proof in a model with both TD and SD components). For evidence supporting this claim see Gagnon (2009) and Alvarez et al. (2015) who show that decision rules are quite insensitive to the inflation for rates that are below 10 per cent.

[7] 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).

friction is irrelevant (provided the economies have the same frequency and kurtosis of price changes).

The explanation of why this result holds is involved, but its interpretation is not. The ratio in equation (3) controls for both the degree of flexibility of the economy, as measured by $N(\Delta p_i)$, as well as for the presence of "selection" effects, as measured by $Kurt(\Delta p_i)$.[8] On the one hand, that the cumulative impulse response depends on the degree of flexibility is hardly surprising. On the other hand, that the selection effect is captured completely by the steady state kurtosis of prices is, at least to us, more surprising. Moreover, that exactly the same expression holds for state dependent and time dependent models is, again at least to us, revealing. In summary, the reason is that the selection effect operates equally in terms of the size distribution of price changes (which is the mechanism for state dependent models) as well as on the distribution of times between adjustments (which is the mechanism for time dependent models). Our result states that as long as any two models produce the same level of kurtosis (of the size of price changes) as well as the same average frequency of price changes then the total cumulated output response produced by a monetary shock is the same across these models, in spite of the fact that their underlying frictions might differ.

The third theoretical result is that TD and SD models behave differently in response to large shocks. Using the notation of equation (1) we have that $\Theta'(\delta) = 0$ for any value of the shock δ in TD models. This is intuitive since the timing of pricing decisions is independent of the state by definition. Thus TD models imply an impulse response function for the aggregate price level is proportional to the size of the shock. Formally for all shock sizes δ there is no impact effect on prices in TD models, so that $\Theta(\delta) = 0$. Moreover, TD models have a proportional flow effect $\theta(\delta, t) = \theta(1, t)\,\delta$ at all horizons $t \geq 0$. These two results imply that

$$\text{TD models: } \mathcal{P}(\delta, t) = \mathcal{P}(1, t)\,\delta\,, \text{ for all } \delta$$

so that the function $\mathcal{P}(\cdot, t)$ is linear with a zero intercept.

Instead, in SD models we have that $\Theta'(0) = 0$ but $\Theta'(\delta) > 0$ for $\delta > 0$, and thus $\Theta''(0) > 0$. In particular SD models imply a minimum shock size such that all shocks above this size give rise to full price flexibility (monetary neutrality). Formally, we can show that there is a shock $\bar{\delta} < \infty$ such that for all $\delta \geq \bar{\delta}$ we have $\Theta(\delta) = \delta$ and $\theta(\delta, t) = 0$ or, that the economy displays full price flexibility for sufficiently large shocks. Thus in SD models $\mathcal{P}(\cdot, t)$ has unit

[8]The selection effect, a terminology introduced by Golosov and Lucas (2007), indicates that firms that change prices after the monetary shock are the firms whose prices are in greatest need of adjustment, a hallmark of SD models.

derivative with respect to δ for large values of δ:

$$\text{SD models: } \mathcal{P}(\delta,t)=\delta \text{ for all } t \text{ if } \delta \geq \bar{\delta} \quad , \quad \text{otherwise} \quad \mathcal{P}(\delta,0)=\frac{1}{2}\Theta''(0)\delta^2+o(\delta^2) \quad .$$

2.1 Overview of empirical analysis

This section presents a preliminary investigation of the non-linear response to nominal shocks that was discussed in the theoretical part. We follow the ideas in Burstein and Gopinath (2014) and Campa and Goldberg (2005), as well many others in the pass-through literature, and use nominal exchange rate fluctuations as a proxy for an "orthogonal" nominal shock to the firms' nominal costs. Since we seek to identify the different behavior of the economy conditioning on the *size* of the exchange rate shocks, it is important that we have a large number of observations to be able to include as many episodes as possible of small as well as of large shocks.

We use an unbalanced panel of monthly data from the post-Bretton woods period for about 70 countries in periods of moderate inflation. Our focus on moderate inflation countries is suggested by the theory: as inflation increases the impact effect $\Theta(\delta)$ of a small nominal shock δ becomes larger, so that the difference between a small and a large shock becomes harder to detect.[9] Since the data are monthly we cannot really estimate the impact effect $\Theta(\delta)$, but we can measure the CPI change after the shock $\mathcal{P}(\delta,t)$ where t, the time elapsed, is 1 month.[10] The monthly data provide yet another reason to focus on low inflation: while the frequency of price adjustment is unresponsive to inflation at low inflation rates, the frequency increases as inflation enters the 2 digit range (see the evidence in Gagnon (2009) and Alvarez et al. (2015)), so that the propagation of shocks is faster and its shape becomes harder to detect.

We compute an inflation forecast at different horizons conditional on an exchange rate innovation, using a simple non-linear regression.[11] We measure the pass-through from exchange rate changes to inflation for $t = 1,3,6,12$ and 24 months, allowing for the magnitude of the pass-through to depend on the size of the exchange rate shock. A key issue in estimation concerns the simultaneous interaction between inflation and the nominal exchange rate,

[9] Alternative definitions of moderate inflation are used: our baseline requires that the mean inflation rate is below $X\%$ in a 10 year time window centered on the observation date. Our baseline results use $X = 8$ but results are robust to using a threshold of $X = 6$. Smaller thresholds reduce the number of large devaluations observed in sample. The mean unconditional annual inflation in our baseline sample is below 4%, see the summary statistics reported in Table 1.

[10] This is a marginal improvement of the early analysis of Caballero and Engel (1993) who measured the inflation response during the year after the shock.

[11] Our baseline specification uses simple non-linear projections as suggested in Jorda (2005), but results are robust to the distributed lag specification commonly used in the international economics literature.

which opens the possibility of reverse causation. We think reverse causation is especially likely for countries in a fixed exchange rate regime, where large devaluations may occur as a "realignment" after periods of above average inflation. For this reason we also control for the type of *de facto* exchange rate regime distinguishing between flexible, managed and fixed exchange rates (as classified by Levy-Yeyati and Sturzenegger (2003); Reinhart and Rogoff (2004) and the following update in Ilzetzki, Reinhart, and Rogoff (2008)). Also, the (near) random-walk nature of exchange rates in floating-exchange rate countries makes such a sample a more appropriate to use our specification to test the hypothesis of the differential (in terms of size) impact effect of exchange rate shocks.

We test whether the short-term pass through, namely the conditional correlation between the nominal exchange rate innovations and inflation (at various horizons), is bigger for large exchange rate movements than for small ones. This is because the theory of SD models predicts a larger response of inflation to nominal shocks in the presence of large shocks, while TD models predict the shape of the impulse response function to be independent of the size of the shock. Various non-linear functional forms were considered: a quadratic specification, a cubic as well as a piecewise linear specification. In Table 2 we report the estimates of the quadratic specification:

$$\pi_{i,(t,t+h)} = \alpha_i + \delta_t + \beta_h \, \Delta e_{i,t} + \gamma_h \, (\Delta e_{i,t})^2 \; sign \, (\Delta e_{i,t}) + \epsilon_{it}^{\pi} \tag{4}$$

where $\pi_{i,(t,t+h)}$ is the inflation rate of country i in the period from month t to month $t+h$ (for $h = 1, 3, 6, 12, 24$), $\Delta e_{i,t}$ is the devaluation from month $t-1$ to month t.[12] The sign operator is used to impose "symmetry", i.e. that the inflation effect of a large devaluation equals the deflation effect of a large appreciation. All regressions use time and country dummies (fixed effects) and standard errors are computed using STATA's robust standard error options (similar results obtain by clustering errors at the period or country level).

Our empirical results, summarized in Table 2, uncover some evidence of a non-linear effect, i.e. of a statistically significant γ_h coefficient, in the sample that excludes countries in a fixed exchange rate regime.[13] The top panel of the table shows that the impulse response of inflation to large shock is above the response to a small shock up to the 6-month horizon. After 24 months the two impulse responses coincide, with a pass through of about 10%. This

[12] To be precise, devaluation is computed as $\Delta e_{i,t} = (e_{i,t}/e_{i,t-1} - 1) \times 100$ where $e_{i,t}$ is the end of the period bilateral exchange rate of country i against the US and inflation is computed as $\pi_{i,(t,t+h)} = (p_{i,t+h}/p_{i,t} - 1) \times 100$ where $h = 1, 3, 6, 12, 24$ months and $p_{i,t}$ is the price level reported for period t. Note that the CPI $p_{i,t}$ is constructed using prices that are sampled *during* period t; that is, between the end of period $t-1$ and the end of period t.

[13] This result is robust to alternative classifications of the de facto ER regime, such as Ilzetzki, Reinhart, and Rogoff (2008), used in the table, vs the one by Levy-Yeyati and Sturzenegger (2003), which includes fewer countries, and was used by us for a robustness check.

is consistent with the hypothesis that larger nominal shocks have a shorter half-life. This effect is quite robust in the sample of flexible exchange rate countries (more below). However the results are more noisy when the sample is extended to include the countries in a fixed exchange rate regime.

To quantify the magnitude of the non-linear effect after a shock of size δ we measure the degree of convexity of the estimated impulse response function $\mathcal{P}(\delta, h)$ relative to its linear part given by $\mathcal{P}'(0, h) \cdot \delta$ where h is the 1 month horizon and the prime denotes the derivative with respect to δ. Using the quadratic specification given above we can thus define

$$\Gamma(\delta, 1) \equiv \frac{\mathcal{P}(\delta, 1)}{\mathcal{P}'(0, 1) \cdot \delta} = 1 + \frac{\gamma_1}{\beta_1}\delta \tag{5}$$

It is apparent that for small values of δ the specification is approximately linear (i.e. $\Gamma \approx 1$), a case which is consistent with TD models, as discussed above. But for large shocks $\Gamma(\delta, 1) > 1$ provided that γ_1 is not zero. For instance, using the estimates for the sample that excludes the fixed exchange rate countries we see that a 5.5% shock gives a $\Gamma \approx 1.7$ so that the linear specification underestimates the impact effect (over the first month) by about 70%.[14]

Extensive robustness analysis of the data reveals that the non-linear effects described in the top panel of Table 2 are strongest in the sample of countries classified as either belonging to a flexible or managed exchange rate regime. The non-linear effects are robust to alternative definitions of what constitutes a "low inflation" country (see the robustness analysis in Section 5). The data sample from the Bretton Woods period (1946-1973) confirms that the non-linear effect is not apparent for fixed exchange rate countries. Instead, the non-linear effects appear in the post-1990 sample over all countries (although again it is stronger for low inflation countries in a flexible exchange rate arrangement) and it is also found to hold with different functional form specifications (piecewise-linear, cubic, or distributed lags), different controls (e.g. controls for the fixed vs flex exchange rate regime), as discussed in Section 5.

3 Set up for Model Economies

In this section we describe the class of model economies for which we characterize the effect of a once and for all monetary shock. Section 3.1 defines the economic environment within which firms operate. In Section 3.2, Section 3.3 and Section 3.4 we describe setups where the price setting rule followed by the firms are, respectively, state-dependent, time-dependent and of a mixed type. In these problems the firms take a constant interest rate as well the common

[14]The shock $\Delta e = 5.5$ is equal to 2 standard deviations of the measured percent changes of Δe, see Table 1.

part of their nominal marginal cost as given. Our main results are about the response of the aggregate price level to a once and for all common change in the nominal marginal cost. In a sense, this is an "industry analysis" as, for instance, in Eichenbaum, Jaimovich, and Rebelo (2011). In Appendix A we describe a set up where the results hold and can be interpreted as the general equilibrium response to a nominal shock in a closed economy model.

3.1 Firm's price setting problem

We first describe the static production function of the firms, and then we define the price gaps, a concept we will use to characterize the firm's decision rules.

Production Each firm k produces and sells a quantity y_{ki} of n goods (each indexed by i), each with a linear labor-only technology with productivity $1/Z$:

$$y_{ki}(t) = \frac{\ell_{ki}(t)}{Z_{ki}(t)} \quad \text{where} \quad Z_{ki}(t) = \exp\left(\sigma\, \mathcal{W}_{ki}(t)\right)$$

where $\ell_{ki}(t)$ its the labor input. Firm k is subject to a productivity shock that is common across all its products, $\bar{\mathcal{W}}_k$, as well as to idiosyncratic productivity shocks $\tilde{\mathcal{W}}_{ki}$, independent across products. In particular we assume the log of productivity follows a brownian motion $\mathcal{W}_{ki}(t)$ with variance σ^2, namely:

$$\mathcal{W}_{ki}(t) = \frac{\bar{\sigma}}{\sqrt{\bar{\sigma}^2 + \sigma^2}}\, \bar{\mathcal{W}}_k(t) + \frac{\sigma}{\sqrt{\bar{\sigma}^2 + \sigma^2}}\, \tilde{\mathcal{W}}_{ki}(t) \tag{6}$$

where the processes $\{\bar{\mathcal{W}}_k(t), \tilde{\mathcal{W}}_{ki}(t)\}$ are independent across k and i. In words, the process for $\{\mathcal{W}_{ki}\}$ are independent across firms, have a common component with volatility $\bar{\sigma}$, and product specific volatility σ.

Profit function and price gaps. For all the model specifications that we consider we can define a "price gap", g_{ki}, namely the log difference between the current nominal price and the static profit maximizing price for good i sold by firm k. In particular we let $P_{ki}(t)$ be the nominal price at time t and $W(t)Z_{ki}(t)$ be the nominal marginal cost of production for good i and firm k, where $W(t)$ is the nominal wage at time t. Each firm faces a demand with a constant elasticity η for the bundle of its n products, which has an elasticity of substitution ϱ between each of its n varieties.[15] We define the price gap $g_{ki}(t)$ as the log of the difference

[15]See equation (40) and equation (41) in Appendix A for a model where price gaps are derived from primitives.

between the current price and the static profit maximizing price:

$$
\begin{aligned}
g_{ki}(t) &= \log P_{ki}(t) - \log\left(W(t) Z_{ki}(t)\right) - \log\left(\eta/(\eta-1)\right) \\
&= \log P_{ki}(t) - \mathcal{W}_{ki}(t) - \log W(t) - \log\left(\eta/(\eta-1)\right)
\end{aligned}
$$

where we omit the firm k subindex whenever it causes no misunderstanding. Since we consider the case of constant inflation π which induces a constant drift in the nominal wage $W(t)$, and productivity follows a brownian motion, the law of motion of the price gaps will also follow a brownian motion with drift equal to minus the inflation rate and possibly with correlation between products. Thus, absent a price adjustment, each price gap g_i has continuous paths:

$$
dg_{ki}(t) = -\pi\, dt + \sigma d\mathcal{W}_{ki}(t)\,. \tag{7}
$$

We let $\Pi\left(P_{k1}(t), ..., P_{kn}(t), Z_{1k}(t), ..., Z_{kn}(t), W(t); c(t)\right)$ denote the nominal profits of firm k, i.e. its total nominal revenue minus production costs. We can approximate this profit function around the frictionless profit maximizing prices as:

$$
\begin{aligned}
&\Pi\left(P_{k1}(t), ..., P_{kn}(t), Z_{1k}(t), ..., Z_{kn}(t), W(t); c(t)\right) \qquad (8)\\
&= W(t)\left[\frac{\varrho(\eta-1)}{2\,n}\left(\sum_{i=1}^{n} g_{ki}^2(t)\right) - \frac{(\varrho-\eta)(\eta-1)}{2\,n^2}\left(\sum_{i=1}^{n} g_{ki}(t)\right)^2\right] \\
&+ o\left(||(c(t), g_{k1}(t), \ldots, g_{kn}(t))||^2\right) + \text{terms independent of } g_k(t)
\end{aligned}
$$

where $o(x)$ a function of order smaller than x. This second order approximation is useful because it simplifies the objective function to be used in the dynamic problem. Notice that profits can be expressed as a function of price gaps. The variable $c(t)$ stands for any variable that enters in the profit function in a weakly separable way. For instance, in the general equilibrium model of Section A, $c(t)$ corresponds to the aggregate consumption. In that model $c(t)$ is a shifter of the quantity demanded –due to its effect on the aggregate ideal price index– and also indirectly affects the present value of profits through its effect on the real rate. Nevertheless, up to a second order, we argue we can disregard these effects. We make this approximation precise in Section A.1.

3.2 State dependent pricing rules

We describe a price setting problem where the firm's optimal decision rule are *state dependent*, i.e. where price changes occur when the state, given by current profits or markups, attains a critical level. We assume that firms has to pay a fixed *menu cost* to simultaneously adjust the price charge for the n products it produces.

State dependent and Ss decision Rules. We let $g = (g_1, \ldots, g_n)$ be the vector of the n price gaps for the firm, where we omit the firm index k for simplicity. A *state dependent decision rule* is described by an inaction set $\mathcal{I} \subset \mathbb{R}^n$ and a value of the price gap $g^* \in \mathcal{I}$. Given these two elements the optimal state dependent decision rule is inaction if $g(t) \in \mathcal{I}$, and otherwise if $g(t) \notin \mathcal{I}$, then prices are changed so that the vector of price gaps right after the adjustment equal $g(t) = g^*$. We note that if $g^* = 0$, i.e. if the n price gaps are set to zero, then it means that when prices are adjusted they are all set to a value that maximizes the static profits.

In general the inaction set can be described by a function $b : \mathbb{R}^n \to \mathbb{R}$:

$$(g_1, \ldots, g_n) \in \mathcal{I} \implies b(g_1, \ldots, g_n) \le 0 \text{ and } (g_1, \ldots, g_n) \notin \mathcal{I} \implies b(g_1, \ldots, g_n) > 0 \tag{9}$$

We will consider the case where the n products enter symmetrically, so that b is symmetric and the n elements of g^* are identical. Symmetry provides a convenient mapping from the inaction set $\mathcal{I}$ into an Ss rule given by two (scalar) threshold functions, one for the lower bound $\underline{g} : \mathbb{R}^{n-1} \to \mathbb{R}$, and one for the upper bound $\bar{g} : \mathbb{R}^{n-1} \to \mathbb{R}$. In an Ss rule prices are changed if, given the rest of the price gaps, the price gap of any product (say product one) reaches either a lower threshold $\underline{g}$, or an upper threshold $\bar{g}$, i.e.:

$$(g_1, g_2, \ldots, g_n) \in \mathcal{I} \iff \underline{g}(g_2, \ldots, g_n) \le g_1 \le \bar{g}(g_2, \ldots, g_n) \tag{10}$$

Summarizing, we can describe an Ss rule by the optimal return point g^* and either a function b or the pair of functions $(\underline{g}, \bar{g})$.

Micro-foundation of state dependent model. We can microfound the state dependent rules described in equation (9) or equation (10) as the solution of the following problem. Consider a firm that chooses when to change prices, i.e. the stopping times $\{\tau_i\}$ as well as the price changes $\Delta P_j(\tau_i)$ at those times to maximize:

$$\begin{aligned} \max_{\{\tau_i, \Delta P_j(\tau_i), j=1,\ldots,n, i=1,2,\ldots\}} \mathbb{E} \Bigg[\int_0^\infty & e^{-rt}\, \Pi\left(\{P_1(t), ..., P_n(t), Z_1(t), ..., Z_n(t)\}\,, W(t)\right)\, dt \\ & - \sum_{i=1}^{\infty} e^{-r\tau_i}\, \psi_m\, W(\tau_i) \Bigg] \end{aligned} \tag{11}$$

$$P_j(t) = P_j(\tau_i) \;\text{ for all } t \in (\tau_i, \tau_{i+1}] \;\text{ and }\; \Delta P_j(\tau_i) = \lim_{\epsilon \downarrow 0} P_j(\tau_i + \epsilon) - P_j(\tau_i)$$

The two main parameters for this class of models are the size of the menu cost ψ_m and the number of products n. The key assumption for the multi product specification (where $n > 1$) is that once the menu cost ψ_m is paid the firm can adjust the prices of all goods at no extra cost. We will provide an analytic characterization of this non-concave stochastic sequence problem by solving an approximate version which uses the quadratic profit function defined in equation (8). Several models discussed in the recent literature are nested as special cases of the state dependent setup. We briefly recall some of them next.

Classic menu cost. The menu cost problem, as in Golosov and Lucas (2007), is obtained setting $n = 1$. In this model the menu cost is constant at ψ and with zero inflation the optimal policy is the well known Ss rule: firms adjust their prices when the distance (in absolute value) between the actual price and the profit maximizing price gap reaches a value $\pm\bar{g}$. This model produces a size-distribution of price changes that is degenerate: when the price adjustments occur and are of size $\pm\bar{g}$.

Multiproduct models. This model allows for $n \geq 2$ and any value of π. The setup with $n = 2$ and normal innovations to productivity gives Midrigan (2011), with $n = 3$ gives Bhattarai and Schoenle (2014), and for large values of n (technically $n \to \infty$, but in practice for $n > 10$) the model produces Taylor's (1980) staggered model, where the time elapsed between two price adjustments is constant. Alvarez and Lippi (2014) show that with zero inflation $\pi = 0$, and with $\eta = \varrho$ (which implies that the elasticity of substitution between bundles is the same as the elasticity between varieties in a bundle), then the function b describing the optimal Ss rule can be written as:

$$b(g_1, \ldots, g_n) = \sum_{i=1}^{n} g_i^2 - \bar{y} \text{ and } g_i^* = 0 \text{ for all } i = 1, \ldots, n\,, \tag{12}$$

for an optimally determined value of $\bar{y}$. Equivalently we can write b in terms of the optimal thresholds:

$$\underline{g}(g_2, \ldots, g_n) = -\left(\bar{y} - \sum_{i=2}^{n} g_i^2\right)^{1/2} \quad \text{and} \quad \bar{g}(g_2, \ldots, g_n) = \left(\bar{y} - \sum_{i=2}^{n} g_i^2\right)^{1/2} \tag{13}$$

The key economic insight of this model is that this framework generates small price changes, since the stopping times (for price adjustments) are defined by the sum of n price gaps, which implies that an individual price gap at the time of adjustment can take any value in

$(-\sqrt{\bar{y}}, \sqrt{\bar{y}})$.[16]

In the case where the elasticity of substitution $\eta \neq \varrho$, and/or there is steady state inflation, so $\pi \neq 0$, and/or there is correlation between the idiosyncratic shocks to the products, so $\bar{\sigma} > 0$, we have that the function b that defines the set of inaction $\mathcal{I}$ can be written as:

$$b(g_1, \ldots, g_n) = \sum_{i=1}^{n} g_i^2 - \bar{y}\left(\sum_{i=1}^{n} g_i\right) \tag{14}$$

where, with a slight abuse of notation, we use $\bar{y}$ to denote a function $\bar{y} : \mathbb{R} \to \mathbb{R}$. In this more general case we can define two scalars $y \equiv \sum_{i=1}^{n} g_i^2$ and $z \equiv \sum_{i=1}^{n} g_i$ which we can use to define the inaction set. Moreover, one can show that the diffusions for (y, z) follow themselves a first order Markov process, i.e. so they are sufficient to define the state of the problem.[17]

3.3 Time Dependent pricing rules

We describe a price setting problem where the firm's optimal decision rule are *time dependent*, i.e. where the time between consecutive price changes is statistically independent of the current markup or profits of the firm. Under such a rule the time elapsed since the last price change completely determines the hazard rate of price changes.

More formally, let an observation be an event in which the firm collects and process all the information that is necessary for price setting. Absent other frictions, observation time will coincide with the times of a change of prices, to adapt to the newly gathered information. Let τ_i be the date of the i^{th} observation: at this time the firm uses all available information to adjust its price(s) and to decide the time of the next observation, τ_{i+1}. Formally we allow for random dates in the sense that $\tau_{i+1} - \tau_i$ is random variable with (right) cumulative distribution function H, i.e $\Pr\{\tau_{i+1} - \tau_i \geq t \,|\, \tau_i\} = H(t|\, \tau_i)$. The defining characteristic of a time dependent model is that the realization of τ_{i+1} is independent of the information relevant for price setting, i.e it is independent of the price gaps $\{g_1(t), \ldots, g_n(t)\}$ for $t \geq \tau_i$. Note that with this definition price changes cannot have any selection, where we use selection in the sense of Golosov and Lucas (2007).

Well known examples of TD models are Taylor (1980) model of staggered price setting, where price adjustments are deterministically spaced every T periods, or $H(t) = 1$ for $t < T$, and $H(t) = 0$ otherwise. Another well known example is the model by Mankiw and Reis (2002) where the times elapsed between successive observations are exponentially distributed,

[16] See Alvarez and Lippi (2014) for an analytical characterization of the optimal stopping barrier $\bar{y}$, as well as the implication for the size distribution of price changes $f(\Delta p_i)$.

[17] See Section 6 and Appendix E of Alvarez and Lippi (2014) for a proof.

or $H(t) = e^{-\lambda t}$ so that the mean time elapsed between observations is $1/\lambda$.[18] More general versions of these models allow the distribution of times to follow a first order Markov process $H(t;t_0)$ where the distribution of times elapsed between observations t is allowed to depend on duration of the previous spell between observations t_0.

Aggregating the behavior across firms, each described by the function H, provides a characterization of the stationary cross-sectional distribution of the "times until the next observation": $Q(t)$. That is, the fraction of firms that, at any point in time, will wait at least t units of time until the next observation. We denote the right CDF of such distribution by $Q(t)$, which determines the time it takes for an aggregate shock to be incorporated into the information set of a given fraction of firms, i.e. the speed at which the monetary shock propagates into the aggregate price level.

Micro-foundation of time dependent pricing rules. While the the distribution function H is assumed as a primitive of the analysis in several TD models, the literature following Caballero (1989); Reis (2006) has provided an explicit profit maximization problem subject to information frictions to rationalize the origins of H. We describe a model with explicit microfoundations, based on Alvarez, Lippi, and Paciello (2016), that rationalizes inattentive behavior as the optimal policy given the cost of collecting and processing information. The firm price setting problem balances the costs and benefits of gathering information. We assume that to gather information about the nominal marginal cost the firm must pay an "observation cost", along the lines discussed by Caballero (1989); Reis (2006). In particular, we assume that by paying an observation cost ψ_o, firms learn the current value of the production cost $(Z_1, ..., Z_n)$, which is the key variable to decide prices. We interpret the observation cost as the physical cost of acquiring the information needed to make the price decision as well as costs associated with the decision making in the firm (gathering and aggregating information, e.g. Zbaracki et al. (2004), Reis (2006)). Alternatively, these costs represent the cognitive costs associated with gathering extra information, as found in experimental evidence on tracking problems –see Magnani, Gorry, and Oprea (2016).

The problem for the firm consists in deciding, at each observation date τ_i, the time

[18] To be more precise in the Mankiw and Reis (2002) model prices will change every period to keep up with the mean expected marginal cost. This gives rise to a very high frequency of price changes, that diverges as the model moves to continuous time. This feature is a common element in models of rational inattentiveness that lack a physical cost of price adjustment. A robust pattern in the data is, however, that prices change infrequently. A simple way to obtain infrequent price changes in this class of models is to assume that the level of the nominal marginal cost is a martingale. As a result, price changes only occur when new information arrives, so that the frequency of price changes coincide with the frequency of observations. Moreover, in Alvarez, Lippi, and Paciello (2011) we show that price plans would not be optimal even in the presence of a drift in the nominal marginal cost, when a price adjustment cost is added to a similar model and calibrated to match the frequency of price changes in the U.S. economy.

until the new planned observation date T_i, as well the prices consistent with the available information:

$$\max_{\{\mathrm{T}_i, P_j(t), j=1,\ldots,n, i=1,2,\ldots, t\geq 0\}} \mathbb{E}\left[\int_0^\infty e^{-r\,t}\,\Pi\left(P_1(t),...,P_n(t),Z_1(t),...,Z_n(t)\,,\,W(t)\,;\,c(t)\,\right)\,dt \right.$$
$$\left. -\sum_{i=1}^{\infty} e^{-r\tau_i}\,\psi_o(\tau_i)W(\tau_i)\right] \tag{15}$$

where $\tau_{i+1} = \min\{s_i\,,\,\mathrm{T}_i\} + \tau_i$ where s_i is an exponential distributed r.v. and T_i and $P_j(t)$ for $t \in [\tau_i, \tau_{i+1})$ only depend on information gathered at $\tau_0, \tau_1, \ldots, \tau_i$.

The value of the state of the firm is $(Z_1, \ldots, Z_n)$, which is the information required to set the the prices that maximize current profits. We assume that the state is only observed infrequently. In particular, we assume that there are two ways in which the firm can observe it. First, exogenously and at an exponentially distributed time with duration λ, the state becomes known to the firm. Second, the firm decides when it plans to observe the state. Specifically, we assume that at the time of the i^{th} observation (planned or not), denoted by τ_i the following events take place: (i) the observation cost $\psi_o(\tau_i)$ is realized, (ii) the firm obtains a signal $\zeta(\tau_i)$ that is informative about the value of the next observation cost at different horizons. At this time, the firm decides the time elapsed T_i until then new planned observation, say $\mathrm{T}_i + \tau_i$. Thus the next observation occurs either when the exogenous observation time, denoted by s_i arrives, or when the next planned observation occurs, i.e. $\tau_{i+1} = \min\{s_i\,,\,\mathrm{T}_i\} + \tau_i$. The decision of T_i depends only on the information available at time τ_i. This information consists on the observation costs, signals, and production cost at the current and past observation dates $\tau_0, \tau_1, \ldots, \tau_i$. Furthermore, we assume that (iii) production and observation cost, and exogenous observation times are all statistically independent, and that (iv) nominal marginal cost for each product follows a martingale. Note that (iv) implies that there are no incentives to change prices between observations, i.e. there are no incentives for price plans.[19] Assumptions (i)-(iv) imply that the optimally chosen time between observations, or equivalently between price adjustments, is a function of the signal obtained at the beginning of the price spell, and that the size of price adjustment is independent of the time elapsed between price adjustments, i.e. these assumption imply a *time dependent* model.[20]

Our motivation for introducing the exogenous observation dates, i.e. those triggered by s_i, is to nest the popular model of sticky information where observations (and price changes)

[19]We do this for two reasons. First, when we introduce menu as well as observation costs, price plans will not be optimal for small departures of a martingale–see Proposition 1 in Alvarez, Lippi, and Paciello (2011). Second, if cost are not martingale and there are no menu cost, then price changes will occur as frequently as the model time periods, which will be highly counterfactual.

[20]In Appendix C we give more details on the structure of the cost and signals.

occur with a constant probability per unit of time $\lambda\, dt$. On the other end, by setting $\lambda = 0$, we can abstract from this feature and all the observations involve a cost-benefit analysis in setting τ_i. In general in the determination of τ_i, the value of λ has the same effect as a higher interest rate in the decision of the firm for τ_i.

Our choice of the processes for the observation cost and signals is general enough to nest several cases studied in the literature.

Constant time between observations. If there are no exogenous observation times, i.e. $\lambda = 0$, and observation cost ψ_o are constant, then the time between observation is constant. Caballero (1989) and Reis (2006) analyze this case.

Calvo model. There are two set-ups for this model that give rise to the same distribution of price durations as in the Calvo model. The first one, as explained above, is obtained if all changes are exogenous, i.e. when ψ_o is very large. The second obtains even if $\lambda = 0$ but there is particular distribution of ψ_o and signals so that the firm finds it optimal to observe at exponentially distributed times.[21]

Markovian times. We refer to Markovian times the case where the times between observations (and price changes) form a first order Markov process, so times are random with the current time between observations depending statistically on the duration of the previous spell between observations. This is obtained in a very natural case where the current value of the observation cost is itself the signal for the next observation cost. This case is analyzed both in Reis (2006) (for negligibly small observation costs levels) and Alvarez, Lippi, and Paciello (2016). In particular, we assume that the observation cost follows a continuous time Markov chain, so that for each value θ_o there is a time invariant probability per unit of time to transit to some other values. Thus, when a state is observed by the firm, the value of the observation cost serves as a signal of future observation costs at different horizons, directly implied by the Markov chain. The firm's decision rule becomes a function $\tau(\theta_o)$, so that times between observations are random, each of them corresponding to a value of the realized observation cost θ_o. The economics of this choice balance the benefit of future observations with the expected cost at different horizons. The key property to understand the variability of the τ is the forecastability of future observation cost, which is tightly related

[21] In Alvarez, Lippi, and Paciello (2016), Propositions 7 and 11 show that for any distribution of times, a distribution of signals and cost on future observation costs can be found that provides a foundation to it. This can be used to rationalize the work of researchers that start their analysis directly with the assumption that times between observations (and price changes) are i.i.d. through time with a given distribution, and study their implications for monetary policy, as for example Bonomo, Carvalho, and Garcia (2010).

with the *persistence* of the Markov chain. Furthermore the property of the Markov chain are important to construct the cross sectional distribution of times until the next adjustment $Q(t)$, an object of interest for the impulse response of shocks.

3.4 State and Time Dependent models

We briefly review models that combine state and time dependent elements. In these models the decision rule of the firm depends both on the time elapsed since the last price change as well as on the state (for example on whether markups have reached a critical level). We discuss two examples: the first one is a version of the state dependent model where at exogenously random dates the menu cost is set to zero. The second example is one where there are both observation and menu costs (in Appendix D we write out both problems formally).

Multiproduct Calvo$^+$ model. The first example adds to the state dependent problem described above the arrival of free adjustment opportunities at a constant rate λ. Equivalently, one can interpret the model as one with a random menu cost that, in a period of length dt, equals ψ_m with probability $(1 - \lambda dt)$, or equal zero with probability λdt. This random menu cost introduces exogenous adjustment times in a way that is similar to the one described for the exogenous observation times in time-dependent models. This model is referred to as the Calvo$^+$ model, a combination with features of the Golosov and Lucas model as well as the Calvo model, and was first studied by Nakamura and Steinsson (2010). The multi-product version of this model is studied in Alvarez, Le Bihan, and Lippi (2016).

The optimal policy for this case is a combination of state and time dependent policy. As in the state dependent case, prices change the first time the boundary of the inaction set is reached. But also, as in the time dependent case, the price changes when a free adjustment opportunity arrives. Thus, the stopping time τ_{i+1} is given by the first time (after the adjustment at τ_i) at which the state either reach the boundary of the inaction set, or that an opportunity to adjust at zero cost occurs.

Positive menu and observation costs. The second example combines both observation and menu costs.[22] Observations will happen at discretely separated periods of times, and we denote the i^{th} observation by the stopping time τ_i. Observations are subject to a constant fixed cost $\psi_o > 0$. Upon an observation the firm will decide whether to adjust prices or not,

[22]For simplicity, and because its effects are mainly covered by the Calvo+, we abstract from the exogenous observation times featured in equation (15) for the time dependent case.

which we denote by the indicator $a(\tau_i) \in \{0, 1\}$. If a price adjustment occurs, then a menu costs $\psi_m > 0$ must be paid.

The optimal decision rule combines features of the state and time dependent rules. Upon the current observation at time τ_i and with the relevant information gathered until that time (i.e. the value of the state) the firm decides the time until the next observation, T_i, so that the next observation occurs at $\tau_{i+1} = \tau_i + \mathrm{T}_i$. Also upon an observation at time τ_i, the firms decides whether to adjust prices or not, i.e. whether $a(\tau_i) \in \{1, 0\}$. Note that the pricing decision has a state dependent feature, in that upon an observation the firm will keep prices constant if the state is in the inaction set, and adjust them otherwise. But, differently from the menu cost model described above, upon an observation the firm may find its state strictly outside of the inaction region, i.e. it may strictly prefer to adjust. Versions of this model are analyzed in Alvarez, Lippi, and Paciello (2011, 2016), Bonomo, Carvalho, and Garcia (2010); Bonomo et al. (2016) .

3.5 Steady state price setting statistics

For future reference we introduce three steady state statistics that summarize price setting behavior. We consider a steady state with idiosyncratic shocks and no aggregate shocks and define: i) $N(\Delta p_i; \pi)$ the average number of price changes per unit of time (say per year), $Var(\Delta p_i; \pi)$ the standard deviation of the distribution of the size of (non-zero) price changes and iii) $Kurt(\Delta p_i; \pi)$ the kurtosis of the distribution of the size of (non-zero) price changes. For future reference we index each of these statistics with the steady state inflation rate π. Notice that, in principle, these statistics are directly measurable in micro data sets.

We focus on these statistics because for the three classes of models described above the ratio $Kurt(\Delta p_i; \pi)/N(\Delta p_i; \pi)$ provides a sufficient summary statistics for the real effects of a small monetary shock. This is the object of the next section. In Appendix E we show that there is an identity between $N(\Delta p_i)$ and $Std(\Delta p_i)$ for small inflation rate, uncovering the trade off that any decision rule around zero inflation faces.

4 The propagation of monetary shocks

This section provides an analytic characterization of the propagation of a once and for all nominal shock. The main results are summarized by three propositions that are useful to understand the similarities and the differences between state-dependent and time-dependent models.

We study the effect on aggregate prices of a once and for all shock of size δ on nominal costs. In particular, we start with a pre-shock path of wages $\bar{W}(t) = \bar{W} e^{\pi t}$ for $t \geq 0$ and a pre-shock steady-state equilibrium aggregate price level $\bar{P}(t) = \bar{P} e^{\pi t}$. Following the shock at time $t = 0$ we get $W(t) = e^{\delta}\ \bar{W}(t)$ all $t \geq 0$ and that, under the simplifying assumptions discussed in Appendix A, the nominal interest rate and the nominal wages satisfy

$$R(t) = r + \pi \ , \ \log \frac{W(t)}{\bar{W}(t)} = \delta \text{ for al } t \geq 0 \tag{16}$$

In Appendix A we also give an interpretation of this shock as the general equilibrium response of a closed economy to a once and for all shock to the money supply. Using this model we can also compute the impulse response of output to this shock. In this section we refer to this as the GE version. In this version the pre-shock level of output is constant.

Letting $P(t)$ be the path of prices after the shock, we can write:

$$\begin{aligned} \log \frac{P(t)}{\bar{P}(t)} &= \delta + \int_0^1 \left(\frac{1}{n} \sum_{i=1}^{n} (g_{ki}(t) - \tilde{g}_{ki}) \right) dk \\ &+ \int_0^1 \left(\sum_{i=1}^{n} o(||p_{ki}(t) - \tilde{p}_{ki}(t)||) \right) dk \end{aligned} \tag{17}$$

where $\tilde{g}_{ki}$ are the price gaps in the steady state *before* the shock and $o(x)$ denotes a function of order smaller than x.[23] This expression motivates the study of the impact of the monetary shock on the price level. The expression for the ideal price level is further simplified by noticing that in the case of an steady state with zero inflation, given the approximations derived in Section A.1, price gaps integrate to zero, i.e.:

$$\int_0^1 \left(\frac{1}{n} \sum_{i=1}^{n} \ \tilde{g}_{ki} \right) dk = 0 \ \text{ if } \pi = 0 \tag{18}$$

Next we use the results derived above to study the effect of an aggregate monetary shock of size δ on the aggregate price level $P(t)$ at $t \geq 0$ periods after the shock, which we denote by $\mathcal{P}(\delta, t)$:

$$\mathcal{P}(\delta, t) \equiv \delta + \int_0^1 \left(\frac{1}{n} \sum_{i=1}^{n} g_{ki}(t) - \tilde{g}_{ki} \right) dk = \log \frac{P(t)}{\bar{P}(t)} \tag{19}$$

We emphasize that the time $t = 0$ price gaps g_{ki} are thus right after the shock, so they can change on impact. Once we characterize the effect on the price level, we describe the effect

[23] To understand this equation notice that the monetary shock increases the desired prices of all firms by the same amount δ. This implies a decrease by an amount δ of all price gaps g_{ki} that are not adjusted on impact (the gaps $\tilde{g}_{ki}$ are unaffected by definition). This explains the presence of the term δ in the equation.

on employment and output. The impulse response is made of two parts: an instantaneous impact adjustment (a jump) of the aggregate price level which occurs at the time of the shock, denoted by $\Theta(\delta;\pi)$, and a continuous flow of adjustments from $t>0$ on, denoted by $\theta(\delta,t)$. The impact effect can be written as

$$\Theta(\delta;\pi) = \int_0^1 \left(\frac{1}{n} \sum_{i=1}^{n} \log P_{ki}(0) - \log \tilde{P}_{ki}(0) \right) dk = \log \frac{P(0)}{\bar{P}} \tag{20}$$

where $\tilde{P}_{ik}(0)$ is the steady state price before the shock.[24] Thus cumulative effect of the price level $t \geq 0$ periods after the shock is

$$\mathcal{P}(\delta,t) \;\; = \;\; \Theta(\delta) + \int_0^t \theta(\delta,s)ds \; . \tag{21}$$

We also study the impact effect on the fraction of firms that change prices, denoted by $\Phi(\delta)$.

Notice that

$$\begin{aligned} \Theta(\delta;\pi) &= \delta + \int_0^1 \left(\frac{1}{n} \sum_{i=1}^{n} \left[g_{ki}(0) - \tilde{g}_{ki} \right] \right) dk \\ &= \delta + \int_0^1 \left(\frac{1}{n} \sum_{i=1}^{n} g_{ki}(0) \right) dk \text{ if } \pi = 0 \end{aligned} \tag{22}$$

where $\tilde{g}_{ki}$ are the price gaps just before the monetary shock. Right after the monetary shocks price gaps will change to $g_{ki}(0)$. This change has two parts. First, mechanically, the log nominal wage increases by δ, so that every single price gap decreases by δ. Notice that this mechanical effect cancels with the first term δ on equation (22). Second, as a consequence of the changes in wages some firms may decide to adjust their prices right after the shock occurs (depending on the type of model), so that for those products and firms there is an extra change in the price gap $g_{ki}(0)$. To see this notice that the price gap of firm k product i right after the shock can be decomposed as the price gap pre-shock $\tilde{g}_{ki}$ minus the common increase in wages δ plus the increase in prices $\log P_{ki}(0) - \log \tilde{P}_{ki}(0)$

$$g_{ki}(0) = \tilde{g}_{ki} - \delta + \left(\log P_{ki}(0) - \log \tilde{P}_{ki}(0) \right) \tag{23}$$

The first two terms are mechanical, and the third is the only one that depends on what the firms do. We introduce price gaps, instead of working directly with the price increases, due to two reasons. One is that in state dependent models price gaps are the state, hence it

[24] In the GE version describe in Appendix A this is an approximation of the ideal price index since we are using an equal weighed index.

facilitates to understand what will happen. Second, in all models, the contribution of a firm k product i to the output deviation from steady state can be written in terms of the price gap.

We note a few properties of the impulse response of prices: i) the impact effect is bounded by δ, ii) in the long term the shock is completely pass-through to prices, iii) in the flexible price case prices jump on impact:

$$0 \leq \mathcal{P}(0, \delta; \pi) \equiv \Theta(\delta; \pi) \leq \delta, \; \lim_{t \to \infty} \mathcal{P}(t, \delta; \pi) = \delta, \text{ and } \mathcal{P}^{flex}(t, \delta; \pi) = \delta \,.$$

where we use the super-index $flex$ for the flexible price case.

GE version and effect on output. In our GE version, output and prices are tightly negatively related after the shock, so we can easily compute the output effect. The negative relationship comes from the assumption that agents are on their labor supply schedule, and that nominal wages jump on impact. Thus the effect on output mirrors the one on real wages. This logic gives:

$$\mathcal{Y}(\delta, t; \pi) = \frac{1}{\epsilon} \left[\delta - \mathcal{P}(\delta, t; \pi) \right] \tag{24}$$

where $1/\epsilon$ is a parameter describing the uncompensated labor supply elasticity, as described in Appendix A. We define, as a summary measure of the impulse response, its cumulative version, i.e. the area under equation (24)

$$\mathcal{M}(\delta; \pi) = \int_0^\infty \mathcal{Y}(\delta, t; \pi) \, dt \tag{25}$$

For the cumulative effect of output we also have

$$0 \leq \mathcal{Y}(0, \delta; \pi) \equiv \frac{1}{\epsilon} \left[\delta - \Theta(\delta; \pi) \right] \leq \frac{\delta}{\epsilon} \,, \; \lim_{t \to \infty} \mathcal{Y}(t, \delta; \pi) = 0 \text{ for all } \delta$$
$$\mathcal{Y}^{flex}(t, \delta; \pi) = 0 \text{ for all } t \geq 0 \text{ and thus } \mathcal{M}^{flex}(\delta; \pi) = 0 \text{ for all } \delta \,.$$

4.1 Small shocks: State and time dependent models are identical

In this section we discuss two results: one for the impact and another for the cumulative output (and prices) effect that occur following a small monetary shocks. The main finding is that these effects are the same for all the models considered here, i.e. those models with a general equilibrium set up as described in Section A and with state dependent decision rules -as described in Section 3.2, and/or time dependent decisions rules -as described in Section 3.3, and/or those with features of both -as described in Section 3.4.

The first result in Proposition 1 concerns the impact effect on prices and output of a once and for all monetary shock, as defined in equation (20).[25] The second result is for the GE version of our model, where we use the cumulative output response of an once and for all monetary shock, i.e. the area under the impulse response for output, as defined in equation (25). The result for the impact effect is more general in scope: for instance, it holds for all levels of the inflation rate $\pi = 0$ and holds if $\eta \neq \varrho$, and/or if there is correlation between the idiosyncratic shocks across the products of the firm, i.e. if $\bar{\sigma} > 0$.[26] The second result in Proposition 2 is obtained analytically for a smaller class of economies, and with the GE interpretation of this shock. In particular, it is obtained around a zero inflation rate, and we also restrict the analysis to the case of same elasticity of substitution $\varrho = \eta$, and no correlation between idiosyncratic shocks, i.e. $\bar{\sigma} = 0$.

We start by considering the impact effect of a small shock on prices and output.

Proposition 1. Let $\Theta(\delta;\pi)$, defined in equation (20), be the impact effect on the price level of a once and for all monetary shock of size δ for an economy starting at steady state inflation π. Fix the inflation rate π. If the decision rules are *state dependent* as in equation (10), and $\sigma > 0$, then:

$$\Theta(\delta;\pi) = \Theta'(0;\pi)\,\delta + o(\delta) \quad \text{with} \quad \Theta'(0;\pi) \equiv \frac{\partial}{\partial \delta}\Theta(\delta;\pi)\Big|_{\delta=0} = 0 \tag{26}$$

where $o(\delta)$ means of order smaller than δ. If the decision rules are *time dependent* as in equation (15), then:

$$\Theta(\delta;\pi) = 0 \;\text{ for all } \delta \;\text{ and for all }\; \pi. \tag{27}$$

The proposition states that there is no first order effect of a small monetary shock in either a SD and as well as in a TD model. The result is stronger for time dependent models in the sense that the impact effect is zero for any size of the monetary shock. We give a brief intuitive explanation of the result. Note that for both time and state dependent models, the aggregate shock increases wages by δ log points, and thus decreases every price gap. Firms in time and state dependent models react differently. In both cases the firms that change their prices on impact will change, on average, by a discrete amount proportional

[25] This statistic was first used by Caballero and Engel (1993, 2007) to summarize the degree of flexibility of an economy.

[26] Indeed for the first result we only use the form of the decision rules, either them be state dependent Ss rules as in equation (10), or time dependent rules as in equation (15).

to δ. Nevertheless, in both cases we will conclude that the fraction of firms that adjust on impact, denoted by $I(\delta)$ is of order smaller than δ, i.e. we argue that $I'(0) = 0$. The argument why $I'(0) = 0$ is different for time than for state dependent models.

In the case of state-dependent models, described in Section 3.2, the reason why $I'(0) = 0$ is that the firm's decision to change prices depends on whether the state after the shock is outside the set of inaction $\mathcal{I}$. For instance, in the one product case $n = 1$, prices are adjusted when the price gap reach either boundary of the range of inaction given by an interval $\left[\underline{g}\,,\,\bar{g}\right]$. A key argument in this models is that in steady state, right before the shock occurred, there is a zero density at the boundaries of the range of inaction. Then the fraction of firms that adjust is proportional to δ, for small δ. In particular, denoting by f the steady state density of the price gaps, the fraction that adjust equals $I(\delta) = \int_{\underline{g}}^{\underline{g}+\delta} f(g)dg = f(\underline{g})\,\delta + o(\delta)$. The fact that in steady state there is zero density around the boundary, i.e. that $f(\underline{g}) = f(\bar{g}) = 0$, is a general feature of the state dependent Ss decision rules with idiosyncratic shocks ($\sigma > 0$), and it is so because at the boundary of the range of inaction firms exit (i.e. prices are adjusted) when they get an idiosyncratic shock that will push them outside. Interestingly, this argument extends to multiproduct $n \geq 1$, with correlated shocks across products, and inaction sets given implicitly by ranges equation (10).

In the case of time dependent models, described in Section 3.3, the reason why $I'(0) = 0$ is that the firm's decision rules depend on the time elapsed since the last adjustment. Hence, even if the price gap changes, the firm will not be aware of it until the next review time (decided in the past) comes due. Finally, since the model is set in continuous time, on impact there is a negligible fraction of firms adjusting, i.e. the number that adjust in an interval of length dt equals $(1/N(\Delta p_i; \pi))\,dt$. Hence, as $dt \to 0$, the fraction of firms that change prices to go zero and $I(\delta) \to 0$ for any δ! Note that we are assuming that the information about the change in the price gap due to change on wages is not observed until the time at which firms have previously schedule their decision to learn the state, which is the key assumption of time-dependent models based on inattentiveness. The argument for models with both time and state dependent features, as in Section 3.4, is more complicated, but unsurprisingly the results still holds.

The logic of the general proof . The next paragraphs illustrate the logic of the proof of Proposition 1 for SD models in the general case with many goods and allows for correlated shocks across goods. The proof has two parts. The first part shows that the steady state density f of price gaps g evaluated at the boundary of the inaction set is zero. We write this as Lemma 1 and include its proof in the Appendix B.

LEMMA 1. *Assume $\sigma > 0$ for a SD model. Then, there is zero density at an exit point, i.e.*

if $b(g)=0$, then $f(g)=0$.

The logic of the result in Lemma 1 is easier to see in the one dimensional case ($n=1$), which we present separately. The idea behind this result is that the boundary of the inaction set is an exit point, i.e. if a firm price gap hits the boundary it will change the price, discretely changing the price gap. This behavior, where the steady state mass "escapes" with non-negligible probability to a discretely far away regions of the state space implies that the steady state density has to be zero (this is in contrast with the behavior everywhere else where the mass moves only to closed-by states). Specifically, consider a discrete state discrete space. We let the time periods be of length Δ and a state space for the price gap be of size $\sqrt{\Delta}\sigma$. The price gap $g(t+\Delta)-g(t)=\sqrt{\Delta}\sigma$ with probability $\frac{1}{2}\left[1-\frac{\pi\sqrt{\Delta}}{\sigma}\right]$ and down to $-\sqrt{\Delta}\sigma$ with the complementary probability. Thus the expected change and expected square change of g per period are $-\pi\Delta$ and σ^{Δ} respectively. The range of inaction is given by an interval $\left[\underline{g},\bar{g}\right]$. We write the analog to the Kolmogorov forward equation in discrete time for the probability of each value g in the state space as for any $g\neq g^*$:

$$f\left(g;\Delta\right)=\begin{cases} f\left(g-\sqrt{\Delta}\sigma;\Delta\right)\frac{1}{2}\left[1-\frac{\pi\sqrt{\Delta}}{\sigma}\right] & \text{for } g\leq\bar{g} \\ f\left(g-\sqrt{\Delta}\sigma;\Delta\right)\frac{1}{2}\left[1-\frac{\pi\sqrt{\Delta}}{\sigma}\right]+f\left(g+\sqrt{\Delta}\sigma;\Delta\right)\frac{1}{2}\left[1+\frac{\pi\sqrt{\Delta}}{\sigma}\right] & \text{for } \underline{g}\leq g\leq\bar{g} \\ f\left(g+\sqrt{\Delta}\sigma;\Delta\right)\frac{1}{2}\left[1+\frac{\pi\sqrt{\Delta}}{\sigma}\right] & \text{for } g\geq\underline{g} \end{cases} \tag{28}$$

At the upper bound we have:

$$f(\bar{g})=\lim_{\Delta\downarrow}f\left(\bar{g};\Delta\right)=\lim_{\Delta\downarrow}f\left(\bar{g}-\sqrt{\Delta}\sigma;\Delta\right)\lim_{\Delta\downarrow}\frac{1}{2}\left[1-\frac{\pi\sqrt{\Delta}}{\sigma}\right]=f(\bar{g})\frac{1}{2} \tag{29}$$

where we use that, provided that $\sigma>0$, thedensity $f(\cdot)$ is continuous in the closure of the range on inaction in the first and last equalities. We obtain that the only possible solution of $f(\bar{g})=f(\bar{g})/2$ is $f(\bar{g})=0$. An analogous argument shows that $f(\underline{g})=0$.

The second part shows that the impact effect on aggregate prices is of second order with respect to δ. In the general case we define the fraction of firms (or price gap vectors) that adjust prices in impact as $I(\delta)$ as

$$I(\delta)=\int_{-\infty}^{\infty}\left[\cdots\int_{-\infty}^{\infty}f(g_1,g_2,\ldots,g_n)1_{\{b(g_1-\delta,g_2-\delta,\ldots,g_n-\delta)>0\}}dg_n\cdots\right]dg_1 \tag{30}$$

where we use that $f(g)=0$ if $b(g)<0$. Thus $I(\delta)$ integrates using the density f the firms whose price gaps will be outside the set of inaction, i.e. $b(g_1-\delta,g_2-\delta,\ldots,g_n-\delta)>0$, after the aggregate shock δ. We set the second part as Lemma 2:

Lemma 2. Assume that there is no density on the boundary of the inaction set. Then, there is no first order impact effect on prices, i.e. $I'(0) = 0$.

In the one dimensional case Lemma 2 follows from direct computation of I and of its derivative, as shown in the text. The main idea is that the firms that change prices on impact for a small aggregate shock δ are those close to the lower boundary of the inaction set, since price gaps decrease all by δ. Thus if the density of price gaps are zero at the lower boundary, there is no first order effect on the fraction of firms adjusting. The $n-$dimensional case is more involved, in this case the price gap for each firm is a vector for which each of its components decreases by δ with the shock. In the $n > 1$ case we also has to take into account a general (unknown) shape of the n-dimensional set of inaction and the correlation among the price gap from different products of the firm. In particular, in the $n > 1$ case there is no simple lower bound for the range on inaction as is is in the one dimensional case. We prove Lemma 2 by finding a function $\bar{I}(\delta)$ which is a suitable upper bound for $I(\delta)$. The upper bound function $\bar{I}(\delta)$, which is inspired by the one dimensional, also has zero derivative when evaluated at $\delta = 0$. While technically the proof is more involved, the logic is the same as in the one dimensional case: for a small aggregate shock δ the firms that will change its price in impact has to belong to the boundary of the set of inaction.

How the impact effect varies with inflation. We conclude the analysis of the impact effect with a discussion of the role of inflation, π, relative to the volatility of the idiosyncratic shocks σ. Let us focus on a positive monetary shock $\delta > 0$ in the reminder of this paragraph. This shock increases the desired price of all firms by an amount δ. First we note that the impact effect is independent of the inflation rate in TD models just because, by assumption, decision rules do not depend on the state. But inflation does change the the impact effect in SD models. In particular, while for finite values of the ratio π/σ the impact effect Θ is of order δ^2, as stated in equation (27), we notice that the impact effect is increasing with π and that the effect becomes first order as $\pi/\sigma \to \infty$. This is the case, for instance, in the classic menu-cost models of Sheshinski and Weiss (1983) and Caplin and Spulber (1987): in both models the impact effect $\Theta(\delta)$ is of order δ, since in these models $\sigma = 0$ and $\pi > 0$, so that the ratio diverges. Thus, since the impact effect is second order but it is increasing in π, in the empirical analysis we will focus on low-inflation countries where the lack of response to small shocks should be easier to detect.

We briefly expand on the reason why in the case of $n = 1$, as π/σ increases, the impact effect of an aggregate monetary shock increases. As explained above, the first order term on $\Theta(\delta, \pi)$ is given by the invariant density $f(\underline{g}, \pi)$. A straightforward analysis of the Kolmogorov forward equation solved by f shows that as inflation rises relative to the variance of the idiosyncratic shock σ^2, the shape of $f(\cdot,\pi)$ changes in the segment $[\underline{g}, g^*]$ as follows. The density f is linear in g for $\pi = 0$, and it becomes concave in g for $\pi > 0$, with curvature $-f''(g,\pi)/f'(g,\pi) = 2\pi/\sigma^2$. In the limit as $\pi/\sigma^2 \to \infty$ the density $f(\underline{g})$ is strictly positive, and $f(\cdot)$ is constant, so that there is a first order effect. Note that this is the case in the classical analysis of Sheshinski and Weiss (1983) and Caplin and Spulber (1987), because there are no idiosyncraric shocks $\sigma = 0$. The reason why the invariant distribution "piles up" more density around $\underline{g}$ as inflation rises is straightforward: the price gaps drift to $\underline{g}$ at speed π, and they only go up when they are hit by a positive idiosyncratic shock (with variance σ^2).

In Appendix F we formally analyze how the impact effect varies with inflation around small inflation rates. For simplicity we focus on a model with one good $n = 1$ and assume the adjustment thresholds $\bar{g}, \underline{g}$ and optimal return point g^* are fixed at the level corresponding to zero inflation.[27] Since $f(\underline{g};\pi) = 0$, expanding the first non-zero term of the impact effect Θ as a function of the inflation rate to obtain:

$$\begin{aligned}\Theta(\delta;\pi) &= \frac{1}{2} f'(\underline{g};\pi)\delta^2 + o(\delta) \approx \frac{1}{2}\left[f'(\underline{g};0) + f'_\pi(\underline{g};0)\pi\right]\delta^2 = \frac{1}{2}\left[\frac{1}{\bar{g}^2} + \frac{1}{\sigma^2\bar{g}}\pi\right]\delta^2 \\ &= \frac{1}{Std\left[\Delta p_i\right]}\left[\frac{2}{Std\left[\Delta p_i\right]} + \frac{\pi}{\sigma^2}\right]\delta^2\end{aligned}$$

The approximation shows that the impact effect is increasing in the ratio π/σ^2, since around zero inflation the steady state standard deviation of price changes is given by $Std\left[\Delta p_i\right] = 2\bar{g}$.

Finally, we note that in both TD and SD models a higher inflation rate tends to increase the average number of price adjustments per unit of time $N(\Delta p_i;\pi)$, even though the elasticity of the frequency of price adjustment to the inflation rate is zero at $\pi = 0$ in models with $\sigma > 0$ (see Alvarez, Lippi, and Paciello (2011); Alvarez, Le Bihan, and Lippi (2016) for formal proofs). This implies that for small rates of inflation (or deflation) the frequency of price adjustments is very close to the frequency that occurs at zero inflation (see Alvarez and Lippi (2014) for a formal proof). In practice, this theoretical prediction is consistent with evidence on the small elasticity of the frequency of price changes in Gagnon (2009) and Alvarez et al. (2015), who show that the frequency is basically insensitive to inflation for rates between 5 and 10 % (in absolute value).

[27] In Proposition 3 of Alvarez et al. (2015) we show analytically for the same model that around zero inflation only 10% of the changes in inflation are accounted for changes in the thresholds, and instead 9/10 are accounted for by changes in the frequency of price increases vs price decreases. This means that ignoring the changes in thresholds, as in the approximation above, makes a very small difference.

The output effect in a simple GE model. Next we turn to characterize the cumulative output effect $\mathcal{M}(\delta)$ of a shock δ. We show that this effect is well approximated by the ratio of two steady state statistics: $N(\Delta p_i;\pi)$, the average number of price adjustments per unit of time, and $Kurt(\Delta p_i;\pi)$, the kurtosis of the size distribution of (non-zero) price changes. These two statistics, in turn, depend on all the structural parameters of the model. But once the models for which the proposition applies are matched with these two statistics, then these models will have the same cumulative effect after a small monetary shock.

Proposition 2. Let $\mathcal{M}(\delta;\pi)$, defined in equation (25), be the cumulative impulse response of output to a once and for all monetary shock of size δ for an economy starting at steady state inflation π. Then

$$\mathcal{M}(\delta;\pi) = \frac{Kurt(\Delta p_i;0)}{\epsilon\, 6\, N(\Delta p_i;0)}\,\delta + o\left(||(\delta,\pi)||^2\right) \tag{31}$$

where $o(x)$ means of order smaller than x. Moreover,

$$\frac{\partial}{\partial\pi}\left(\frac{Kurt(\Delta p_i;\pi)}{N(\Delta p_i;\pi)}\right)\Bigg|_{\pi=0} = 0 \tag{32}$$

The explanation of why this result holds is involved, but its interpretation is not. The ratio in equation (31) controls for both the selection effect, as measured by $Kurt(\Delta p_i;\pi)$, and for the degree of flexibility of the economy, as measured by $N(\Delta p_i;\pi)$. On the one hand, that the cumulative impulse response depends on the degree of flexibility is hardly surprising. On the other hand, that the selection effect is captured completely by the steady state kurtosis of prices is, at least to us, more surprising. The role of kurtosis is more novel and embodies the extent to which "selection" in the size as well as in the timing of price changes occurs.[28] The selection effect, a terminology introduced by Golosov and Lucas (2007), indicates that firms that change prices after the monetary shock are the firms whose prices are in greatest need of adjustment, a hallmark of SD models. Selection gives rise to large price adjustments after the shock, so that the CPI response is fast.[29] Such selection is absent in TD models where the adjusting firms are chosen based on (possibly stochastic

[28]For a symmetric distribution kurtosis is a scale-free statistic describing its peakedness: the extent to which "large" and "small" observations (in absolute value) appear relative to intermediate values.

[29] Intuitively, a lot of selection gives rise to small kurtosis. For example, in the Golosov-Lucas model 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 in a multi product model with a large number of goods is normally distributed, i.e. it features a large mass of small as well as very large price changes. This results in less selection, fully captured by the higher kurtosis of the size distribution.

functions of) calendar time, not based on their state. In addition to selection in the *size* of price changes, recent contributions have highlighted a related selection effect in TD models which relates to the *timing* of price changes.[30] Surprisingly, the kurtosis of the steady-state distribution of the size of price changes also encodes this type of selection, which is central to TD models. 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.[31] This paper collects and extends previous results by showing that equation (31) also holds for models with both TD and SD components. Formally, we proved the result for SD models in Alvarez, Le Bihan, and Lippi (2016) and the result for TD models in Alvarez, Lippi, and Paciello (2016) for the case of $n = 1$ products.[32] For the multiproduct version of the Calvo$^+$ model, which showcases features of both SD and TD model, the result is also shown in Alvarez, Le Bihan, and Lippi (2016).[33] In Appendix G we provide numerical evidence that the result also holds in models that combine those frictions. In that appendix we illustrate this result by aggregating and computing the impulse response of a decisions rules for a price setting problem with both a menu cost as well as an observation cost, based on Alvarez, Lippi, and Paciello (2011, 2015, 2016).[34]

Extending the expression for the area under the impulse response for TD models to n products. Here we argue that Proposition 1 and 2 in Alvarez, Lippi, and Paciello (2016) hold with no changes for the $n > 1$ case, which establishes Proposition 2 for the TD case. Consider the TD model in its multiproduct version. In Alvarez, Lippi, and Paciello (2016) we show for that model with $n = 1$ that equation (31) holds. This results extends in a straightforward way to the multiproduct case of $n > 1$. To see why, let τ be the time elapsed between observations, noticing that since the menu cost is zero, τ it is also the time

[30]See Kiley (2002); Sheedy (2010); Carvalho and Schwartzman (2015); Alvarez, Lippi, and Paciello (2016).

[31] 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).

[32]To be precise for SD models, it follows by setting $\lambda = 0$, or equivalently $\ell = 0$, in the model in Alvarez, Le Bihan, and Lippi (2016), and the results corresponds to Proposition 6 of that paper.

[33]Again, this corresponds to Proposition 6 of that paper, for the case where $\lambda > 0$ and $\psi > 0$, or equivalently the case where $\ell \in (0, 1]$

[34]These computations take advantage that we have a characterization of the decision rules that allows to compute a simple problem for each ratio of the menu to observation cost, free of any other parameters.

between price changes. Recall that when an observation occurs, every single product of the firm change its price, "closing" its gaps (this, instead, extend from the one to the n products since it only requires the symmetry or exchangeability, and the lack of drift). Thus the state of the economy is still the distribution of times until the next review, the same as the one dimensional object that in Alvarez, Lippi, and Paciello (2016) is denoted by $Q(t)$ with density $q(t)$. At the time of the adjustment (or of the observation) we can then consider each of the n products of the firms in isolation. This is because the marginal distribution of the price gaps of each of the n products is the same, and hence the result is identical. Note that this results holds even if the the price gaps have an arbitrary correlation between the products of the same firm. It only requires that the marginal distribution of each price gaps be normal.

The expression in equation (31) can be regarded as a *second order* approximation to $\mathcal{M}(\delta;\pi)$. This expression means that

$$0 = \frac{\partial \mathcal{M}(0,0)}{\partial \pi} = \frac{\partial^2 \mathcal{M}(0,0)}{\partial \delta^2} = \frac{\partial^2 \mathcal{M}(0,0)}{\partial \pi^2} = \frac{\partial^2 \mathcal{M}(0,0)}{\partial \delta \partial \pi} \tag{33}$$

i.e. the approximation in equation (31) holds up to second order, and thus it is very accurate for small values of δ and π. There are two different arguments why these derivatives are zero. First, relative to π, note that by definition $\mathcal{M}(0,\pi) = 0$ for all π, since when there is no shock there is no response. Thus, all derivatives with respect to π are zero at $\delta = 0$. The reason why the second derivatives, especially the one with respect to δ are zero is due to the symmetry of the $\mathcal{M}$ function. In particular $\mathcal{M}(\delta;\pi) = -\mathcal{M}(-\delta;-\pi)$. This means that the effect of of prices and output when there is a negative shock in an economy with deflation is the same (in absolute) value than an economy with inflation and a positive shock. Thus taking any second derivative of this function, and evaluating at $(\delta,\pi) = (0,0)$ we obtain the desired result. Thus, the key is to argue the symmetry of this function. This, in turns, depends, among other things, on the use of the second order approximation of the profit function, as developed in equation (42) to argue for the symmetry of the optimal decision rules. Finally we explain the significance of the fact that the approximation itself has zero derivative with respect to inflation, i.e. the importance of equation (32). This means that the expression for $\mathcal{M}(\delta;\pi)$ is accurate for economies with low inflation rates.

In principle, micro data on prices can be used to construct empirical measures of kurtosis. In constructing such measures care must be taken of small measurement errors a (lots of small price changes are just noise) and heterogeneity (pooling together goods with different volatility of price changes) which may mechanically contribute to generating a high value of kurtosis, as stressed in Cavallo and Rigobon (2016). In Section 2 of Alvarez, Le Bihan, and Lippi (2016) we have used such statistical procedures and estimated Kurtosis values in the neighborhood of 4. This is useful to decide "where the data stand" between a Golosov-Lucas model (with kurtosis 1) vs a Calvo model (with Kurtosis 6).

Three examples. To illustrate the point that models with different degree of time and state dependence can generate the same cumulative output response after a permanent shock we describe three set-ups that give the same value of $\mathcal{M}$ for small δ in spite of their different nature and steady state behavior in other dimensions than those involved by the formula in equation (31). We concentrate on describing $Kurt(\Delta p_i; 0)$, since in all these models it is easy to change other parameters, such as the fixed adjustment or observation cost, to produce the same value of $N(\Delta p_i; 0)$. We focus on three examples where $Kurt(\Delta p_i) = 3$. The first example is a state dependent model with many products, i.e. with $n \to \infty$. This model produces a size distribution of price changes that is Normal (see Alvarez and Lippi (2014) for a proof), so that the kurtosis equals 3. The second example is a pure time dependent model, with constant observation cost $\psi_o >$ (and zero menu cost $\psi_m = 0$). This model is analyzed by e.g. Reis (2006), like the previous model it also produces a size distribution of price changes that is normal, so its kurtosis equals to 3. While the first two models have identical steady state statistics in terms of distribution of adjustment times and the size distribution of price changes, the third one is different. The third example is the so called Calvo-plus model of Nakamura and Steinsson (2010). In this model $n = 1$ and while some prices occur upon the arrival of a free adjustment opportunities, other are decided by the firm after paying the menu cost. Alvarez, Le Bihan, and Lippi (2016) show that if the fraction of price adjustment due to free adjustment opportunities is 90% then the model produces a kurtosis of the size of price changes that is equal to 3 (although the distribution function of the size of price changes is not Normal). Notice that these 3 models are setup to have the same mean duration between price adjustments, but other moments of the distribution of adjustment times will differ. Moreover, the models also differ in terms of the nature of the friction (menu cost vs observation). In spite of these differences, Proposition 2 states that the cumulative output effect of a small monetary shock is identical in these models.

4.2 Large shocks: State and time dependent models differ

In this section we examine the impact effect on prices of large shocks. The result differ between time and state dependent models. For time dependent models, the size of the monetary shock δ is immaterial. Instead for state dependent models, large shocks behave differently than small shocks.

Proposition 3. Consider the impact effect of a once and for all change in money of size

δ for an economy at steady state with inflation rate π. Then in a time dependent model as in Section 3.3 we have:

$$\Theta(\delta;\pi) = 0 \text{ for all } \delta \geq 0\,, \tag{34}$$

while in a state dependent model as in Section 3.2 with $\eta = \varrho$ and no correlation between the idiosyncratic shocks across the products of the firm ($\bar{\sigma} > 0$) we have:

$$\frac{\partial \Theta(\delta;0)}{\partial \delta} \geq 0 \text{ with } \frac{\Theta(\delta;0)}{\delta} \to 1 \text{ as } \delta \to 2\,Std(\Delta p_i;0) \tag{35}$$

and in the general *state dependent* decision rules as in equation (10) with $\sigma > 0$, for each π there is a $\bar{\delta}(\pi)$ such that:

$$\Theta(\delta;\pi) = \delta \text{ for all } \delta \geq \bar{\delta}(\pi)\,. \tag{36}$$

The explanation why the effect of monetary shocks in time dependent models is independent of the size of the monetary shock δ, is familiar from the Calvo model, and it is the exactly the same as the one given for small shocks. The fact that for state dependent models the impact effect is different for large vs small shocks is the hallmark of fixed cost adjustment models. Put it simply, when the shock is large enough, a large fraction of firms will pay the fixed cost and adjust. Interestingly, equation (35) gives a hint of when a shock is large enough so that all the firms will adjust immediately, namely when the shock δ is larger than the steady state standard deviation of price changes. This result can be easily be seen in the case of $n = 1$ product, since the standard deviation of price changes $Std(\Delta p_i) = \bar{g}$, since price changes are $\pm\bar{g}$ with the same probability. Recall that in this case the distribution of price gaps in the steady state right before the shock lies in the interval $[-\bar{g}, \bar{g}]$. Thus, when the shock is large enough so that $\delta > 2\bar{g}$ then every single firm will find that right after the shock has its price gap outside the range of inaction. A similar reasoning holds for any number of products, i.e. for $n \geq 1$. The proof of the result in equation (35), as well as a characterization of this function is given in Proposition 8 (iii) and Proposition 10 of Alvarez and Lippi (2014).

In the general case, fixing all the parameters that define the set of inaction, one can find a value of δ that is large enough so that every price gap vector after the aggregate shock is outside the set of inaction, i.e. $(g_1 - \delta, g_2 - \delta, \ldots, g_n - \delta) \notin \mathcal{I}$. This only requires that the set of inaction $\mathcal{I} \subset \mathbb{R}^n$ is bounded. Thus, one can take $\bar{\delta}(\pi)$ to be the difference between the largest and smallest values in $\mathcal{I}$.[35]

[35] We can take $\bar{\delta}(\pi) = g(\pi)_{max} - g(\pi)_{min}$ where $g(\pi)_{max} = \{\inf x : (x, \ldots, x) \geq (g_1, \ldots, g_n) \text{ for all } g \in \mathcal{I}(\pi)\}$ and $g(\pi)_{min} = \{\sup x : (x, \ldots, x) \leq (g_1, \ldots, g_n) \text{ for all } g \in \mathcal{I}(\pi)\}$.

In Alvarez, Le Bihan, and Lippi (2016) we characterize $\bar{\delta}(0)$, the smallest value of the aggregate shock δ for which all firms adjust their prices on impact for a multiproduct Calvo$^+$ model. In this case for each value of $n \geq 1$ the threshold $\bar{\delta}(0)$ is a function of $\ell \in [0, 1]$, the fraction of all price changes that occur due to the Calvo parameters. As ℓ increases, the value of $\bar{\delta}(0)$ also increases, and indeed as $\ell \to 1$ then $\bar{\delta} \to \infty$. This is quite intuitive: as the importance of the time dependence of the decision rules increase (i.e. as ℓ increases), then threshold for the aggregate shock $\bar{\delta}$ increases.

5 Some evidence

In this section we exploit the predictions of Propositions 1 and 3 to explore the source of the underlying friction. On the one hand, with time dependent rules we expect that the impact effect is independent of the size of the shock. On the other hand, with state dependent rules we expect that the impact effect is second order for small shocks and first order for large shocks. Thus, if the impact of a cost shock on prices depends on the size of the shock, the evidence will point towards state dependence. In the empirical exploration, in particular, we study whether changes on the exchange rate of different sizes imply a differential effect for inflation at different horizons after the shock. We focus on low inflation countries since the approximation of Propositions 1 and 3 is accurate, as discussed in Section 4, for low levels of inflation. In addition, we study the period post Bretton Woods, so that changes in the exchange rate better approximate an unexpected and permanent shock in costs. Overall, in our preliminary results, we find some evidence of non-linear effects. This evidence is stronger for flexible exchange countries and is robust to different regression specifications and definitions of a low inflation country.

Data. We start from the whole sample of Consumer Price Index and Exchange rate data from the International Financial Statistics database from the IMF. For the CPI index we use "CPI of all items". For the Exchange rate we use the end of period exchange rate in units of domestic currency per unit of US dollars.[36] With this data we construct an initial unbalanced panel $\{\pi_{i,t}, \Delta e_{i,t}\}_{i \in \mathcal{I}, t \in T_i}$ where $\mathcal{I}$ is the set of all countries and T_i is the set of dates

[36] In terms of the data by choosing CPI and Exchange rate against the US and working at a monthly frequency we take an alternative route from the literature that studies the exchange rate pass through; see for example Campa and Goldberg (2005), Bussiere (2013) and the Handbook Chapter Burstein and Gopinath (2014). The reasons for studying CPI and the exchange rate against the US Dollar are to obtain as many observations as possible (monthly time series for import prices and effective exchange rate are available only for a subset of countries), and because the model outlined in Section 3 is better suited for the pricing decisions of retailers. In addition, we work at a monthly frequency to better approximate impact effects. It is worth noting that by using Import prices and effective exchange rates we still find some weak evidence of non-linearity. This evidence is in line with Bussiere (2013).

for which observations are available for country i. To be consistent with the setups described in Section 3 and Section 4 we restrict the sample in two dimensions. First, by focusing on low inflation countries. As discussed in Section 4, the result that the impact effect is second order for state dependent models, Proposition 1, is accurate for low levels of inflation because as inflation increases the higher order terms also increase. Still, to identify the effects of large shocks, large devaluations/revaluations are needed in the sample and these events are sometimes associated with countries experiencing moderate and high inflation rates. With this trade-off in mind we restrict the sample as follows: we include the inflation rate of country i in period t in the sample if the 10 year moving average of annual inflation is less than 8 percent (for our baseline specification).[37] Second, we further restrict the sample by focusing on the observations after Bretton Woods. The once and for all monetary shock has two main features: it is unexpected and permanent. The evidence in this direction favors flexible exchange rate countries. To classify a country as a flexible exchange rate country we follow the classifications of Reinhart and Rogoff (2004), Ilzetzki, Reinhart, and Rogoff (2008) and Levy-Yeyati and Sturzenegger (2003). With these two restrictions, we obtain a (unbalanced) panel for our main specification.[38]

Table 1 and Table 7 describe inflation, devaluation, and some features of our panel data. There are 13,025 observations in the main sample. The panel is unbalanced because of different data availability among countries for the CPI and Exchange rate data and because countries enter and exit the sample over time depending on their inflation rates. For inflation, we report the mean and volatility of annual inflation. For devaluation, we report mean and volatility of our main independent variable, monthly devaluation. Note that the mean devaluation is not zero but the mean is usually at least an order of magnitude smaller than the volatility. The number of devaluations/revaluations that are higher than 7, 10 and 15 percent are 368, 131 and 22, respectively. Our preferred specification focuses on the sample of countries that are not classified as fixed exchange rate regimes by Ilzetzki, Reinhart, and Rogoff (2008). For this sample, mean and volatility of inflation is slightly higher and the number of devaluations/revaluations that are higher than 7, 10 and 15 percent are 229, 88, 18, respectively.

[37]With monthly data this amounts to the following restriction. An observation for country i in period t is in the sample if $\pi_{i,t}^{MA} = (\sum_{k=-60}^{k=60} \pi_{i,t+k})/120 \leq \mathcal{K} = 0.08$. We check robustness of our findings for different windows (24, 36, 48 months) and inflation values (4, 6, 8, and 10 percent).

[38]We also focus on countries that have a GDP per-capita higher than 5000 USD whenever the value of GDP per-capita is available. We use the World Bank national accounts data, with data available after 1990 on a monthly basis. Also, we focus on countries that have populations that are higher than 2 million inhabitants.

Table 1: Descriptive Statistics

Sample	Mean(π)	sd(π)	Mean(Δe)	sd(Δe)	# Large innovations $\lvert\Delta e\rvert$		
					>7%	>10%	>15%
Post-1974 Sample, Inflation Threshold 8%							
All countries (13,025 obs)	3.51	3.76	0.08	2.81	368	131	22
No Fixed ER (6,137 obs)	3.14	3.38	0.014	3.00	229	88	18
Post-1990 Sample, Inflation Threshold 8%							
All countries (8,488 obs)	2.94	2.80	0.12	2.86	272	109	18
No Fixed ER (5,010 obs)	2.76	2.75	0.19	3.07	204	82	16

Inflation π is the 12-month percentage change of the CPI. The innovations $|\Delta e|$ are the percent depreciation (or appreciation) of the bilateral nominal exchange rate vs the US dollar over a 1 month period. The criterion for including a country-month observation in the sample is that the 60-month moving average inflation in that month is below 8% (per year) and a per-capita GDP in that country-month of at least \$5,000 (PPP).

Specification. Our baseline specification is given by:

$$\pi_{i,(t,t+h)} = \alpha_i + \delta_t + \beta_h \, \Delta e_{i,t} + \gamma_h \, (\Delta e_{i,t})^2 \;\; sign \, (\Delta e_{i,t}) + \epsilon^{\pi}_{it} \tag{37}$$

where $\pi_{i,(t,t+h)}$ is the inflation rate of country i on the period from date t to date $t+h$, Δe_t is the devaluation from date $t-1$ to date t, and both variables are measured in percent, so that $\Delta e_{i,t} = 1$ is one percent.[39] The structural innovation is given by ϵ^{π}_{it}. The first term in the regression is a fixed effect for country i that captures unobserved effects that are constant over time (for example, the average inflation rate). The second term is a time fixed effect that captures aggregate shocks that are common to the whole group of countries. The third term is the linear component of the pass through where the coefficient β_h measures the impact of a devaluation on period t over inflation on the period that goes from t to $t+h$. The fourth term measures whether large changes in the exchange rate have a higher pass through. If this is the case for horizon h we should expect that $\gamma_h > 0$. Note that the $sign(\cdot)$ operator

[39] To be precise, devaluation is computed as $\Delta e_{i,t} = (e_{i,t}/e_{i,t-1} - 1) \times 100$ where $e_{i,t}$ is the end of the period bilateral exchange rate of country i against the US and inflation is computed as $\pi_{i,(t,t+h)} = (p_{i,t+h}/p_{i,t} - 1) \times 100$ where $h = 1, 3, 6, 12, 24$ months and $p_{i,t}$ is the price level reported for period t. Note that the CPI $p_{i,t}$ is constructed using prices that are sampled *during* period t; that is, between the end of period $t-1$ and the end of period t.

is introduced for symmetry.[40]

One note of caution is due: the regression coefficient can be interpreted as a measure of the response of inflation to an exogenous nominal exchange rate innovation under the assumption that the shock is orthogonal to the other regressors and unanticipated. This assumption, which gave us a motive to focus on flexible exchange rate countries where exchange rates are close to random walks, must be taken with caution. First, the specification implies that changes in inflation do not feed-back in devaluation by directly assuming that the nominal exchange rate follows a Random Walk with its own structural shocks. Second, it can be the case that large swings on the exchange rate are associated with some particular observable or unobservable economic conditions that are not modeled. For example, a large devaluation might occur after a sustained appreciation of the real exchange rate. In this case, a large devaluation could imply a lower pass through (see for example, Burstein, Eichenbaum, and Rebelo (2005) and Burstein and Gopinath (2014)). Another example is that large devaluations occur during bad times, as in Kehoe and Ruhl (2009).

Main Results. The results for our main specification are in Table 2. Overall, we find some evidence of non-linear effects. In particular, we find a statistically significant correlation between large devaluations/revaluations and higher inflation transmission for the complete sample and for the sample where we restrict to countries not classified as having fixed a exchange rate regime as defined in Ilzetzki, Reinhart, and Rogoff (2008). The top panel of Table 2 reports the results of a Panel regression of equation (37) for the sample excluding fixed exchange rate countries post 1974. As one would expect the total pass through of exchange rate into prices increases with the horizon; from 0.01 after one month to around 0.1 after two years for a 1 percent shock. The nonlinear component of the pass through is statistically different from zero, and it is quantitatively relevant for large shocks as discussed in the overview (see equation (5) and the discussion around it). The estimated coefficients imply that a devaluation (revaluation) of 10 percent is associated with a 0.2 percent point

[40]This specification differs from the ones usually estimated in the literature that studies exchange rate pass through (see for example, Campa and Goldberg (2005), and the Handbook Chapter by Burstein and Gopinath (2014)) in two dimensions. First, instead of estimating equation (39) without a non-linear term (that is, $\gamma_h = 0$) several papers estimates a distributed lag regression

$$\pi_{i,t} = \alpha_i + \delta_t + \sum_{s=0}^{T} \beta_s \Delta e_{i,t-s} + \epsilon_{it}^{\pi} \tag{38}$$

It can be shown that under the identifying assumption that the exchange rate follows random walk the two specifications, distributed lags vs linear projections, are analogous. Second, our specification in equation (37) introduces a non-linear term. An exception is Bussiere (2013). This paper runs cross country and country by country non-linear specifications and finds non-linearities in the pass through of the effective exchange rate into import and export prices.

Table 2: Inflation Pass Through: Baseline Specification

	1974-2014 Sample, excluding Fixed ER countries (6,811 obs.)				
horizon h:	1 month	3 month	6 month	12 month	24 month
β_h (linear term)	0.009** (0.004)	0.027*** (0.008)	0.056*** (0.010)	0.053*** (0.016)	0.098*** (0.023)
$\gamma_h \times 100$ (quadratic term)	0.114*** (0.027)	0.152*** (0.054)	0.104 (0.106)	0.111 (0.133)	-0.060 (0.158)
R^2	0.20	0.31	0.41	0.51	0.62

	1974-2014 Sample, All countries (13,273 obs)				
horizon h:	1 month	3 month	6 month	12 month	24 month
β_h (linear term)	0.019*** (0.004)	0.039*** (0.010)	0.062*** (0.013)	0.091*** (0.021)	0.184*** (0.024)
$\gamma_h \times 100$ (quadratic term)	0.058*** (0.028)	0.166 (0.112)	0.097 (0.140)	0.104 (0.257)	-0.448*** (0.149)
R^2	0.11	0.19	0.28	0.39	0.46

	1990-2014 Sample, All countries (9,179 obs)				
horizon h:	1 month	3 month	6 month	12 month	24 month
β_h (linear term)	0.009*** (0.003)	0.033*** (0.006)	0.060*** (0.009)	0.074*** (0.012)	0.109*** (0.019)
$\gamma_h \times 100$ (quadratic term)	0.088*** (0.021)	0.105*** (0.045)	0.008 (0.066)	-0.008 (0.073)	-0.314*** (0.107)
R^2	0.17	0.28	0.37	0.47	0.53

All regressions include time and country fixed effects. Exchange rates for all countries except the US are expressed as the bilateral exchange rate with the US, and as the effective exchange rate for the US. The sample excluding fixed ER countries drops countries with a pre-announced or de facto peg, crawling peg, or band narrower than $\pm 2\%$ using the Ilzetzki, Reinhart, and Rogoff (2008) exchange rate regime classification. One, two or three stars denote the coefficient is statistically different from zero at the 10, 5, or 1% confidence level. The criterion for including a country-month observation in the sample is that the 60-month moving average inflation in that month is below 8% (per year) and a per-capita GDP in that country-month of at least \$5,000 (PPP). Robust standard errors in parenthesis. See Section 5 for details.

of increase in the inflation rate on impact. In addition to the one 1 month pass through, the non-linear component is significant at the 3 and 6 months, horizons. Second, in the mid panel, we report the estimates of equation (37) obtained when using all countries (i.e. not tossing those classified as fixed exchange rate regimes). In this case, the non-linear component decreases but it is still statistically significant at the 10 percent level of confidence. In this case, the total pass through is higher at every horizon compared to the specification without excluding fixed exchange rate countries, both through the linear and non-linear component. The inclusion in the sample of countries which are on a fixed exchange rate arrangement also gives rise, across several specifications, to a significant and negative coefficient for the non-linear term at the 24-month horizon. None of the theories that we reviewed can fit this pattern which we find puzzling. We conjecture that this may be related to the low-pass through of large devaluations: in countries on a fixed exchange rate regime those can happen as a response to a persistent misalignment of the real exchange rate, as documented in Burstein, Eichenbaum, and Rebelo (2005); Burstein and Gopinath (2014). Finally, in the bottom panel of Table 2 we restrict to a sample with both fixed and flexible exchange rate countries but with observations after 1990. The non-linear component is again significant, with a smaller pass through than in the middle panel. This is consistent with the evidence of a lower pass through post 1990 discussed in Taylor (2000).

Robustness. We perform 4 robustness checks to our main specification and sample. We still find some evidence for a non linear pass through. We first run the main specification using a piecewise linear function for the non-linear term. The regression specification is:

$$\pi_{i,(t,t+h)} = \alpha_i + \delta_t + \beta_h \Delta e_{i,t} + \gamma_h \Delta e_{i,t} \mathcal{I}(|\Delta e_{i,t}| > K) + \epsilon^{\pi}_{it} \tag{39}$$

where the only difference with equation (37) is the introduction of $\mathcal{I}(|\Delta e_t| > K)$, as an indicator of whether the devaluation (or revaluation) in period t was higher than K percent in absolute value instead of the quadratic term. We report results for $K = 10$ but we also check robustness for $K = 5, 20$. The results are summarized Table 3. There is evidence of non-linearity for the sample of countries that are not classified as a fixed exchange rate regime. The evidence is weaker for the sample of all countries. The linear portion of the pass-through is in line with the one for the main specification in Table 2. Second, we run the baseline specification under different samples depending on the inflation threshold that we use for the moving average. Recall that in the main sample an observation for country i in period t is in the sample if $\pi^{MA}_{i,t} = (\sum_{k=-60}^{k=60} \pi_{i,t+k})/120 \leq \mathcal{K} = 8$. Results are summarized in Table 4. We find that, if the threshold used is too low, for example a $\mathcal{K}$=4, the non linear term γ_h is not significant. This is also the case if the inflation threshold is set too high. For example for a threshold of 10 percent, results are not significant for the post 1974 samples, and significant at 10 per cent confidence level for the post 1990 sample. Still, the non-linear term cannot be rejected for the thresholds of 5, 6 and 8 percent for the samples post 1974 and post 1990 for all countries.

Finally, we provide two additional robustness tests. First, we exclude from the sample countries that are not classified by Ilzetzki, Reinhart, and Rogoff (2008). This will decrease the number of total observations. Results are in Table 5. In this case, the 2 year linear pass through is again in line with the the main specification. Non-linear effects are rejected for the sample of all countries (fixed and flex), but cannot be rejected for the sample excluding the fixed exchange rate countries, as we can see from the top panel. For the sample of fixed and flexible exchange rate countries post 1990 non-linear effects cannot be rejected. Second, we re-run the main specification in the main sample using the classification in Levy-Yeyati and Sturzenegger (2003). Results are summarized in Table 6. The linear component of the pass through is again similar to the one in Table 2. Non-linear effects cannot be rejected on impact.

Table 3: Inflation Pass Through: Piecewise Linear Specification

	1974-2014 Sample, excluding Fixed ER countries (6,816 obs.)				
horizon h:	1 month	3 month	6 month	12 month	24 month
β_h (linear term)	0.015***	0.034**	0.058***	0.058***	0.0943***
	(0.004)	(0.007)	(0.009)	(0.014)	(0.020)
$\gamma_h \times 100$ (non-linear term)	0.0174***	0.027**	0.027	0.019	-0.005
	(0.008)	(0.016)	(0.023)	(0.035)	(0.049)
R^2	0.30	0.45	0.41	0.51	0.62

	1974-2014 Sample, All countries (13,733 obs)				
horizon h:	1 month	3 month	6 month	12 month	24 month
β_h (linear term)	0.022***	0.047***	0.064***	0.092***	0.157***
	(0.007)	(0.018)	(0.024)	(0.038)	(0.046)
$\gamma_h \times 100$ (non-linear term)	0.012*	0.027	0.029	0.033	-0.047
	(0.007)	(0.018)	(0.024)	(0.038)	(0.046)
R^2	0.16	0.26	0.28	0.39	0.46

	1990-2014 Sample, All countries (9,184 obs)				
horizon h:	1 month	3 month	6 month	12 month	24 month
β_h (linear term)	0.014***	0.038***	0.059***	0.072***	0.100***
	(0.003)	(0.005)	(0.007)	(0.011)	(0.017)
$\gamma_h \times 100$ (non-linear term)	0.014 **	0.017	0.004	-0.020	-0.066**
	(0.006)	(0.011)	(0.014)	(0.022)	(0.032)
R^2	0.30	0.44	0.37	0.47	0.53

All regressions include time and country fixed effects. Exchange rates for all countries except the US are expressed as the bilateral exchange rate with the US, and as the effective exchange rate for the US. The sample excluding fixed ER countries drops countries with a pre-announced or de facto peg, crawling peg, or band narrower than ±2% using the Ilzetzki, Reinhart, and Rogoff (2008) exchange rate regime classification. The piecewise linear specification uses a threshold for large devaluations equal to 10%. One, two or three stars denote the coefficient is statistically different from zero at the 10, 5, or 1% confidence level. The criterion for including a country-month observation in the sample is that the 60-month moving average inflation in that month is below 8% (per year) and a per-capita GDP in that country-month of at least $5,000 (PPP). Robust standard errors in parenthesis. See Section 5 for details.

Table 4: Robustness: Inflation pass through on impact (1 month)

	1974-2014 sample: All Countries				
inflation threshold below:	4%	5%	6%	8%	10%
Non-linear effect	×	✓✓✓	✓✓	✓✓	×
# obs	8,263	9,774	11,030	13,723	16,157

	1974-2014 sample: excluding Fixed ER countries				
inflation threshold below:	4%	5%	6%	8%	10%
Non-linear effect	×	✓✓✓	✓✓✓	✓✓✓	×
# obs	4,566	5,240	5,795	6,811	7,587

	1990-2014 sample: All Countries				
inflation threshold below:	4%	5%	6%	8%	10%
Non-linear effect	×	✓✓✓	✓✓✓	✓✓✓	✓
# obs	6,678	7,651	8,314	9,179	9,813

	1990-2014 sample: excluding Fixed ER countries				
inflation threshold below:	4%	5%	6%	8%	10%
Non-linear effect	×	✓✓✓	✓✓✓	✓✓✓	✓
# obs	4,227	4,751	5,145	5,684	5,997

All regressions include time and country fixed effects. Three, two or one check symbols denote that the coefficient of the non-linear term is statistically different from zero at the 1%, 5% or 10% confidence level (respectively). A cross indicates the coefficient is not statistically different from zero at the 10% confidence level. Standard errors are computed using Stata robust options to deal with minor problems about normality, heteroscedasticity, or some observations that exhibit large residuals, leverage or influence. The sample excluding fixed ER countries drops countries with a pre-announced or de facto peg, crawling peg, or band narrower than ±2% using the Ilzetzki, Reinhart, and Rogoff (2008) exchange rate regime classification. One, two or three stars denote the coefficient is statistically different from zero at the 10, 5, or 1% confidence level. The criterion for including a country-month observation in the sample is that the 60-month moving average inflation in that month is below 8% (per year) and a per-capita GDP in that country-month of at least $5,000 (PPP). Robust standard errors in parenthesis.

Table 5: Robustness: Excluding Unclassified Countries

	1974-2014 Sample, excluding Fixed ER countries (3,896 obs.)				
horizon h:	1 month	3 month	6 month	12 month	24 month
β_h (linear term)	0.008	0.013	0.029**	0.029*	0.081***
	(0.005)	(0.009)	(0.012)	(0.017)	(0.024)
$\gamma_h \times 100$ (quadratic term)	0.096***	0.149***	0.134	0.138	-0.034
	(0.021)	(0.038)	(0.089)	(0.119)	(0.152)
R^2	0.26	0.41	0.55	0.61	0.69

	1974-2014 Sample, All countries (10,808 obs)				
horizon h:	1 month	3 month	6 month	12 month	24 month
β_h (linear term)	0.024***	0.044***	0.062***	0.097***	0.188***
	(0.005)	(0.010)	(0.014)	(0.023)	(0.026)
$\gamma_h \times 100$ (quadratic term)	0.029	0.133	0.079	0.090	-0.434***
	(0.020)	(0.041)	(0.054)	(0.068)	(0.084)
R^2	0.11	0.19	0.29	0.40	0.47

	1990-2014 Sample, All countries (6,436 obs)				
horizon h:	1 month	3 month	6 month	12 month	24 month
β_h (linear term)	0.011***	0.029***	0.045***	0.065***	0.100***
	(0.004)	(0.007)	(0.009)	(0.014)	(0.021)
$\gamma_h \times 100$ (quadratic term)	0.064***	0.082**	0.023	-0.074	-0.294***
	(0.023)	(0.041)	(0.061)	(0.074)	(0.100)
R^2	0.18	0.30	0.42	0.48	0.54

All regressions include time and country fixed effects. We exclude from the sample unclassified countries (Ilzetzki, Reinhart, and Rogoff (2008) de facto exchange rate classification). Exchange rates for all countries except the US are expressed as the bilateral exchange rate with the US, and as the effective exchange rate for the US. The sample excluding fixed ER countries drops countries with a pre-announced or de facto peg, crawling peg, or band narrower than ±2% using the Ilzetzki, Reinhart, and Rogoff (2008) exchange rate regime classification. One, two or three stars denote the coefficient is statistically different from zero at the 10, 5, or 1% confidence level. The criterion for including a country-month observation in the sample is that the 60-month moving average inflation in that month is below 8% (per year) and a per-capita GDP in that country-month of at least $5,000 (PPP). Robust standard errors in parenthesis.One, two or three stars denote the coefficient is statistically different from zero at the 10, 5, or 1% confidence level. The criterion for including a country-month observation in the sample is that the 60-month moving average inflation in that month is below 8% (annualized). Robust standard errors in parenthesis. See Section 5 for details.

Table 6: Inflation Pass Through: Levy-Yeyati and Sturzenegger Classification

	1974-2014 Sample, excluding Fixed ER countries (9,570 obs.)				
horizon h:	1 month	3 month	6 month	12 month	24 month
β_h (linear term)	0.011***	0.031***	0.065***	0.088***	0.125***
	(0.004)	(0.009)	(0.012)	(0.018)	(0.027)
$\gamma_h \times 100$ (quadratic term)	0.108***	0.160**	0.082	-0.010	-0.227*
	(0.036)	(0.076)	(0.121)	(0.133)	(0.130)
R^2	0.29	0.44	0.42	0.54	0.61

	1974-2014 Sample, All countries (13,733 obs)				
horizon h:	1 month	3 month	6 month	12 month	24 month
β_h (linear term)	0.020***	0.040***	0.064***	0.092***	0.184***
	(0.004)	(0.010)	(0.013)	(0.021)	(0.024)
$\gamma_h \times 100$ (quadratic term)	0.057**	0.166	0.091	0.097	-0.448***
	(0.028)	(0.112)	(0.138)	(0.255)	(0.149)
R^2	0.16	0.26	0.28	0.39	0.46

	1990-2014 Sample, All countries (9,184 obs)				
horizon h:	1 month	3 month	6 month	12 month	24 month
β_h (linear term)	0.010***	0.033***	0.060***	0.074***	0.109***
	(0.003)	(0.006)	(0.009)	(0.012)	(0.019)
$\gamma_h \times 100$ (quadratic term)	0.088***	0.109**	0.008	-0.088	-0.314***
	(0.022)	(0.046)	(0.066)	(0.073)	(0.107)
R^2	0.30	0.44	0.37	0.47	0.53

All regressions include time and country fixed effects. Exchange rates for all countries except the US are expressed as the bilateral exchange rate with the US, and as the effective exchange rate for the US. The sample excluding fixed ER countries drops countries with a pre-announced or de facto peg, crawling peg, or band narrower than ±2% using the Levy-Yeyati and Sturzenegger (2003) exchange rate regime classification. One, two or three stars denote the coefficient is statistically different from zero at the 10, 5, or 1% confidence level. The criterion for including a country-month observation in the sample is that the 60-month moving average inflation in that month is below 8% (per year) and a per-capita GDP in that country-month of at least \$5,000 (PPP). Robust standard errors in parenthesis. See Section 5 for details.

Table 7: Countries and main descriptive statistics

	Inflation		Devaluation							
Country	Mean Inf	Sd. Inf	Mean Dev	Sd. Dev	big5	big10	big20	big30	big40	Observations
Algeria	0.78	2.57	0.42	1.80	5	0	0	0	0	179
Austria	0.31	0.71	-0.12	2.64	19	2	0	0	0	503
Belgium	0.33	0.38	-0.04	2.69	23	3	0	0	0	503
Bolivia	4.49	14.38	8.61	81.36	14	13	12	11	10	374
Botswana	0.78	0.69	0.68	3.79	22	8	2	0	0	299
Brazil	5.72	9.16	5.43	9.38	111	56	18	11	3	270
Bulgaria	0.27	0.66	0.24	3.43	8	2	0	0	0	101
Burkina Faso	0.52	3.98	0.10	2.64	17	3	0	0	0	374
Burundi	0.89	2.20	0.45	3.02	6	4	1	1	0	195
Cameroon	0.71	1.72	0.14	3.13	17	3	0	0	0	266
Canada	0.31	0.40	0.05	1.69	6	1	0	0	0	700
Central African Republic	0.35	2.03	0.29	3.58	10	1	0	0	0	108
Chad	0.21	3.77	-0.07	3.49	6	0	0	0	0	77
Chile	0.21	0.38	0.03	3.23	3	1	0	0	0	77
China, P.R.: Hong Kong	0.37	0.77	0.11	0.92	2	0	0	0	0	415
Colombia	1.16	1.11	1.23	5.34	16	5	3	3	2	484
Costa Rica	0.77	1.36	0.63	5.67	14	6	4	3	2	486
Cote d'Ivoire	0.60	2.55	0.11	2.69	17	3	0	0	0	363
Croatia	2.21	7.09	1.96	7.45	33	21	16	4	1	256
Czech Republic	0.28	0.59	0.02	3.69	21	3	0	0	0	245
Denmark	0.39	0.60	0.04	3.01	36	4	0	0	0	581
Dominican Republic	0.72	1.88	0.70	10.70	10	5	1	1	1	460
Egypt	0.72	1.99	0.37	4.50	5	4	2	2	1	399
El Salvador	0.71	1.20	0.25	5.00	1	1	1	1	1	400
Ethiopia	0.53	2.36	-0.06	0.75	0	0	0	0	0	288
Finland	0.50	0.58	0.21	3.28	19	5	2	2	0	503
France	0.47	0.45	0.14	2.93	27	6	0	0	0	503
Germany	0.21	0.33	0.17	3.21	8	1	0	0	0	96
Greece	0.82	1.35	0.51	2.51	24	5	0	0	0	527
Guatemala	0.56	2.04	0.46	7.59	3	2	2	1	1	400
Haiti	0.54	2.79	0.00	0.03	0	0	0	0	0	409
Honduras	0.42	1.01	0.00	0.07	0	0	0	0	0	400
Hungary	0.60	1.21	0.16	3.58	22	8	2	0	0	353
India	0.58	1.14	0.36	3.23	4	1	1	1	1	399
Indonesia	1.07	2.32	0.88	4.89	7	5	3	3	2	269
Iran, Islamic Republic of	0.88	1.60	0.32	4.96	3	1	1	1	1	448
Iraq	0.24	0.93	-0.01	0.05	0	0	0	0	0	51
Ireland	0.14	0.36	0.54	2.63	1	0	0	0	0	24
Israel	1.75	3.44	1.55	5.14	84	32	7	4	3	700
Italy	0.60	0.59	0.22	2.47	26	2	0	0	0	503
Jamaica	0.87	1.26	0.65	4.45	12	5	2	2	2	460
Japan	0.26	0.68	-0.12	2.71	23	2	0	0	0	699
Jordan	0.62	1.95	0.36	2.09	6	1	0	0	0	197
Kazakhstan	0.65	0.56	0.33	2.99	3	2	1	0	0	113
Kenya	0.85	1.21	0.45	2.10	6	3	0	0	0	268
Korea, Republic of	0.57	0.84	0.29	3.45	21	9	3	1	1	545
Kuwait	0.32	0.89	-0.01	0.94	0	0	0	0	0	495
Latvia	1.22	3.75	0.03	3.54	11	2	0	0	0	166
Lebanon	0.30	1.04	0.00	0.00	0	0	0	0	0	81
Libya	0.37	2.57	0.00	1.46	1	1	0	0	0	327
Lithuania	1.48	4.43	0.37	4.42	17	6	3	2	0	272
Macedonia, FYR	0.04	0.43	0.70	2.33	1	0	0	0	0	29
Madagascar	0.87	1.61	0.71	4.77	21	7	2	1	1	315
Malaysia	0.23	0.60	-0.02	1.46	3	0	0	0	0	520
Mali	0.14	1.25	0.14	3.69	4	0	0	0	0	34
Mauritania	0.59	2.64	0.10	2.11	1	0	0	0	0	54
Mexico	1.31	2.05	1.18	6.32	42	10	7	4	4	616
Montserrat	0.21	0.55	0.00	0.00	0	0	0	0	0	170
Morocco	0.46	1.03	0.26	2.34	13	3	0	0	0	399
Myanmar	1.08	2.84	0.05	2.07	4	2	1	1	0	574
Namibia	0.41	0.40	0.83	4.38	9	3	0	0	0	64
Nepal	0.69	2.00	0.44	2.54	5	4	1	1	0	322
Netherlands	0.33	0.69	-0.10	2.68	22	2	0	0	0	503
Niger	0.57	3.05	0.14	3.14	17	3	0	0	0	268
Nigeria	1.02	2.22	0.88	8.04	15	8	3	2	1	365
Norway	0.37	0.56	0.05	2.67	28	4	0	0	0	700
Oman	0.22	0.46	0.00	0.00	0	0	0	0	0	172
Pakistan	0.57	1.33	0.50	6.57	2	1	1	1	1	400
Panama	0.32	0.52	0.00	0.00	0	0	0	0	0	287
Paraguay	0.92	1.98	0.79	6.47	9	5	5	4	3	392

Peru	3.96	8.81	4.24	32.83	60	31	14	9	6	424
Philippines	0.83	1.38	0.70	5.10	11	6	4	2	2	400
Poland	2.11	8.34	2.01	10.26	22	16	7	4	3	184
Portugal	0.83	1.24	0.39	2.66	22	5	1	0	0	503
Qatar	0.15	0.65	0.00	0.00	0	0	0	0	0	94
Romania	2.96	4.67	3.17	14.27	44	19	7	6	4	295
Russian Federation	0.76	0.61	0.61	4.67	12	6	2	0	0	124
Rwanda	0.65	1.83	0.27	6.01	3	2	1	1	1	302
Saudi Arabia	0.13	0.51	0.03	0.19	0	0	0	0	0	423
Senegal	0.64	2.56	0.15	3.13	17	3	0	0	0	269
Serbia, Republic of	0.59	0.82	0.71	4.75	11	4	1	0	0	101
Sierra Leone	0.51		-0.58		0	0	0	0	0	1
Singapore	0.22	0.80	-0.12	1.51	4	0	0	0	0	652
Slovak Republic	0.50	0.86	-0.17	2.87	5	1	0	0	0	156
Slovenia	0.92	1.53	0.71	3.80	18	2	1	1	0	182
South Africa	0.64	0.65	0.45	3.68	40	11	4	0	0	519
Spain	0.70	0.72	0.30	2.98	25	8	2	0	0	503
Sri Lanka	0.57	1.07	0.62	4.80	8	3	2	1	1	399
Sudan	1.31	4.12	0.66	6.75	5	5	4	3	3	387
Sweden	0.36	0.58	0.11	2.80	30	4	0	0	0	700
Switzerland	0.21	0.40	-0.17	3.01	33	4	0	0	0	700
Tanzania	1.04	0.48	1.21	1.47	0	0	0	0	0	3
Thailand	0.44	0.77	0.10	1.33	2	1	0	0	0	353
Togo	0.58	2.52	0.10	3.21	16	2	0	0	0	244
Tunisia	0.53	0.43	0.54	2.28	0	0	0	0	0	47
Turkey	1.91	2.48	1.78	7.07	47	9	5	3	2	389
United Arab Emirates	0.23	0.84	0.00	0.00	0	0	0	0	0	101
United Kingdom	0.22	0.43	0.10	2.80	17	4	0	0	0	329
United States	0.28	0.37	-0.27	1.51	0	0	0	0	0	431
Uruguay	2.72	2.91	2.61	10.17	56	19	7	5	4	532

References

Abel, Andrew B., Janice C. Eberly, and Stavros Panageas. 2007. "Optimal Inattention to the Stock Market." *American Economic Review* 97 (2):244–249.

———. 2013. "Optimal Inattention to the Stock Market With Information Costs and Transactions Costs." *Econometrica* 81 (4):1455–1481.

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 and Francesco Lippi. 2014. "Price Setting With Menu Cost for Multiproduct Firms." *Econometrica* 82 (1):89–135.

Alvarez, Fernando E., Herve Le Bihan, and Francesco Lippi. 2016. "The real effects of monetary shocks in sticky price models: a sufficient statistic approach." *The American Economic Review* forthcoming.

Alvarez, Fernando E., Francesco Lippi, and Luigi Paciello. 2011. "Optimal Price Setting with Observation and Menu Costs." *Quarterly Journal of Economics* 126 (4):1909–1960.

———. 2015. "Phillips curves with observation and menu costs." *Journal of the European Economic Association* forthcoming.

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

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

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

Bonomo, Marco, Carlos Carvalho, Rene Garcia, and Vivian Malta. 2016. "Persistent Monetary Non-neutrality in an Estimated Model with Menu Costs and Partially Costly Information." PUC Rio, Mimeo.

Brunnermeier, Markus K. and Yuliy Sannikov. 2014. "A Macroeconomic Model with a Financial Sector." *American Economic Review* 104 (2):379–421.

Burstein, Ariel, Martin Eichenbaum, and Sergio Rebelo. 2005. "Large Devaluations and the Real Exchange Rate." *Journal of Political Economy* 113 (4):742–784.

Burstein, Ariel and Gita Gopinath. 2014. *International Prices and Exchange Rates*, vol. 4, chap. 7. Elsevier, 391–451. Prepared for the Handbook of International Economics, Vol. IV.

Bussiere, Matthieu. 2013. "Exchange Rate Pass-through to Trade Prices: The Role of Nonlinearities and Asymmetries." *Oxford Bulletin of Economics and Statistics* 75 (5):731–758.

Caballero, Ricardo J. and Eduardo M.R.A. Engel. 1993. "Microeconomic rigidities and aggregate price dynamics." *European Economic Review* 37 (4):697 – 711.

———. 2007. "Price stickiness in Ss models: New interpretations of old results." *Journal of Monetary Economics* 54 (Supplement):100–121.

Caballero, R.J. 1989. "Time Dependent Rules, Aggregate Stickiness And Information Externalities." Discussion Papers 198911, Columbia University.

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

Campa, Jose Manuel and Linda S Goldberg. 2005. "Exchange rate pass-through into import prices." *Review of Economics and Statistics* 87 (4):679–690.

Caplin, Andrew and John Leahy. 1991. "State-Dependent Pricing and the Dynamics of Money and Output." *The Quarterly Journal of Economics* 106 (3):683–708.

Caplin, Andrew S and Daniel F Spulber. 1987. "Menu Costs and the Neutrality of Money." *The Quarterly Journal of Economics* 102 (4):703–25.

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

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

Christiano, Lawrence J., Martin Eichenbaum, and Charles L. Evans. 1999. *Monetary policy shocks: What have we learned and to what end?*, vol. 1 PartA, chap. 2. Elsevier, 65 – 148. Handbook of Macroeconomics.

———. 2005. "Nominal Rigidities and the Dynamic Effects of a Shock to Monetary Policy." *Journal of Political Economy* 113 (1):1–45.

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

Fernandez-Villaverde, Jesus, Pablo Guerron-Quintana, Keith Kuester, and Juan Rubio-Ramirez. 2015. "Fiscal Volatility Shocks and Economic Activity." *American Economic Review* 105 (11):3352–84.

Gagnon, E. 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.

Giavazzi, Francesco, Tullio Jappelli, and Marco Pagano. 2000. "Searching for non-linear effects of fiscal policy: Evidence from industrial and developing countries." *European Economic Review* 44 (7):1259 – 1289.

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

Ilzetzki, Ethan, Carmen Reinhart, and Kenneth Rogoff. 2008. "Exchange rate arrangements entering the 21st century: which anchor will hold?" *University of Maryland and Harvard University* .

Jorda, Oscar. 2005. "Estimation and Inference of Impulse Responses by Local Projections." *American Economic Review* 95 (1):161–182.

Kaplan, Greg and Gianluca Violante. 2014. "A Model of the Consumption Response to Fiscal Stimulus Payments." *Econometrica* 82:1199–1239.

Kehoe, Timothy J and Kim J Ruhl. 2009. "Sudden stops, sectoral reallocations, and the real exchange rate." *Journal of Development Economics* 89 (2):235–249.

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

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-Yeyati, Eduardo and Federico Sturzenegger. 2003. "To Float or to Fix: Evidence on the Impact of Exchange Rate Regimes on Growth." *American Economic Review* 93 (4):1173–1193.

Magnani, Jacopo, Aspen Gorry, and Ryan Oprea. 2016. "Time and State Dependence in an Ss Decision Experiment." *American Economic Journal: Macroeconomics* 8 (1):285–310.

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.

Martins, Fernando. 2005. "The price setting behaviour of Portuguese firms - evidence from survey data ." Working Paper Series 562, European Central Bank.

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

Nakamura, Emi and Jon Steinsson. 2010. "Monetary Non-Neutrality in a Multisector Menu Cost Model." *The Quarterly Journal of Economics* 125 (3):961–1013.

Pollard, Patricia S and Cletus C Coughlin. 2004. "Size matters: asymmetric exchange rate pass-through at the industry level." *University of Nottingham Research Paper* 13 (2004).

Reinhart, Carmen M. and Kenneth S. Rogoff. 2004. "The Modern History of Exchange Rate Arrangements: A Reinterpretation." *The Quarterly Journal of Economics* 119 (1):1–48.

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.

Sheshinski, Eytan and Yoram Weiss. 1983. "Optimum Pricing Policy under Stochastic Inflation." *Review of Economic Studies* 50 (3):513–29.

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

Taylor, John B. 2000. "Low inflation, pass-through, and the pricing power of firms." *European Economic Review* 44 (7):1389–1408.

Woodford, Michael. 2008. "Information-Constrained State-Dependent Pricing." Nber working papers, National Bureau of Economic Research, Inc.

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

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.

A General Equilibrium Set up

This section describes the ingredients of the theoretical results that will be discussed next. Our framework builds on the seminal contributions by Taylor (1980), Calvo (1983), Reis (2006), Golosov and Lucas (2007) and several recent extensions including Midrigan (2011), Nakamura and Steinsson (2010).

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 c, linear in labor hours ℓ, log in real balances M/P, with constant intertemporal elasticity of substitution $1/\epsilon$ for the consumption aggregate, so that the labor supply elasticity to real wages is $1/\epsilon - 1$.

$$\text{HH Lifetime Utility :} \qquad \int_0^\infty e^{-r\,t}\left(\frac{c(t)^{1-\epsilon}-1}{1-\epsilon} - \alpha\,\ell(t) + \log\frac{M(t)}{P(t)}\right)dt \tag{40}$$

$$\text{with CES aggregate :} \qquad c(t) = \left(\int_0^1 \left(\sum_{i=1}^n \left[\,A_{ki}(t)\,c_{ki}(t)\,\right]^{\left(1-\frac{1}{\varrho}\right)}\right)^{\left(\frac{\varrho}{\varrho-1}\right)\left(1-\frac{1}{\eta}\right)} dk\right)^{\frac{\eta}{\eta-1}} \tag{41}$$

The specification assumes a continuum of Dixit-Stiglitz monopolistic sellers, index by k. Each seller sells n goods, indexed by the subscript i, where $\eta > 1$ is the substitution elasticity between sellers (or varieties) and ϱ is the elasticity of substitution between products for each seller. If the elasticities of substitution are the same, i.e. $\varrho = \eta$, then we have the simpler expression:

$$c(t) = \left(\int_0^1 \sum_{i=1}^n \left[\,A_{ki}(t)\,c_{ki}(t)\,\right]^{1-\frac{1}{\eta}} dk\right)^{\frac{\eta}{\eta-1}}$$

To keep the expenditure shares stationary across goods in the face of the permanent idiosyncratic shocks, we assume offsetting preference shocks A_{ki}.[41]

The budget constraint of the representative agent is

$$M(0) + \int_0^\infty \mathcal{Q}(t)\left[\bar{\Pi}(t) + \tau(t) + W(t)\ell(t) - R(t)M(t) - \int_0^1 \sum_{i=1} P_{k,i}(t)c_{ki}(t)dk\right]dt = 0$$

where $R(t)$ is the nominal interest rates, $\mathcal{Q}(t) = \exp\left(-\int_0^t R(s)ds\right)$ the price of a nominal bond, $W(t)$ the nominal wage, $\tau(t)$ the lump sum nominal transfers, and $\bar{\Pi}(t)$ the aggregate (net) nominal profits of firms.

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.

A.1 Optimal Firm Decision Rules and Price Gaps

In this section we show that we, in equilibrium, we can replace the nonlinear profit function of the firm from a simple quadratic problem which depends exclusively on the firm's price gaps. This simplifies the solution of the problem, first by simplifying the state space of the firm, and second by allowing an analytical solution of the firms's decision rules.

[41] This convenient assumption was used by Woodford (2009), Bonomo, Carvalho, and Garcia (2010) Midrigan (2011), Alvarez and Lippi (2014).

We describe how the profits of the firms, once we replace the demand from the household first order conditions problem, as well as using the equilibrium values of nominal wages and nominal interest rates. From here we obtain two results: a description of the state of the firm's problem, and a characterization (up to second order) of the objective function of the firm. We note that firms's profit depend on nominal wages, nominal interest rates, but also aggregate consumption, trough its determination of equilibrium real rates as well as a shifter of the firm's individual demand. Thus if we let $\mathcal{V}(\boldsymbol{p_k}, \mathbf{c}; p_k)$ be the value of the firm k gross profits (i.e. without subtracting the observation and/or menu costs) as a function of initial price gap vector p_k, and for an arbitrary stochastic process for prices $\boldsymbol{p_k}$ and a path of aggregate consumption $\boldsymbol{c}$ we can show that:

$$\begin{aligned} \mathcal{V}(\boldsymbol{p_k}, \mathbf{c}; p_k) \;&=\; -\Upsilon\left(\frac{W(0)}{\bar{W}(0)}\right)\mathbb{E}\left[\int_0^\infty e^{-rt}B\left(\sum_{i=1}^n g_{ik}^2(t)\right)dt \,\middle|\, g_k(0)=g_k\right] \\ &+\; \mathbb{E}\left[\int_0^\infty e^{-rt}o\left(||(g_k(t), c(t)-\bar{c})||^2\right)\,dt \,\middle|\, g_k(0)=g_k\right] + \iota(\delta, \mathbf{c}) \end{aligned} \tag{42}$$

where $\Upsilon > 0$ is a function only on $W(0)/\bar{W}$ and where $\iota(\cdot)$ is only a function of δ and the path of consumption and where $o(x)$ denotes a function that is of smaller order than x. In particular, this means that there are no interactions between $g_{ki}(t)\; c(t)$, and hence $c(t)$ does not impact, up to first order, the determination of the optimal prices, provided that the price gaps and the shock are both small. Moreover, if we include the menu and observation cost, they can be measured in terms of frictionless profits -for instance the normalize menu cost will be $\psi = \psi_m/\hat{\Pi}(0)$ where the normalized profit function $\hat{\Pi}$ is defined below. Importantly, note that aggregate consumption does not feature on this problem, i.e. it does not interact with the price gaps. Finally the constant $B = (1/2)\,\eta\,(\eta-1)$. If the elasticity of substitution η between firms and the n products produced by a firm have elasticity ϱ is different, then instead of $B\left(\sum_{i=1}^n g_{ki}^2(t)\right)$ the quadratic approximation gives:

$$\frac{\varrho(\eta-1)}{2n}\left(\sum_{i=1}^n g_{ki}^2(t)\right) - \frac{(\varrho-\eta)(\eta-1)}{2n^2}\left(\sum_{i=1}^n g_{ki}(t)\right)^2 \tag{43}$$

Note this is a function of two scalars, the sum of the squares of the price gaps, omitting the firm's index k we have: $y \equiv \sum_{j=1}^2 g_j^2$ as well as the sum of price gaps $z \equiv \sum_{j=1}^2 g_j$.

Equation (17) and equation (46) means that to study the effect of a shock on the deviations from output we can study the contribution to the price level of each firm. equation (42) means that to study the optimal price setting each form can regard its objective function to be quadratic and ignore all the other general equilibrium effects. Moreover, since wages adjust on impact, we can simple reset every single firm price gap to be δ smaller log points, and track its aggregate effect keeping the same optimal rules than in steady state.

The remaining of this section provides the details to show the result in equation (42) first, given that cost shocks follow a random walk, and that nominal interest rates are constant, wages growth at a constant rate, then current profits of the firm can be written as a function of price gaps and of exogenous process. Second, consider the discounted nominal profits of the firm k can we written as, where for simplicity we consider the case with $\eta = \varrho$:

$$\begin{aligned} & \mathcal{Q}(t)\, W(t)\, A_{ki}(t)\, Z_{ki}(t)^{1-\eta} \quad \tilde{\Pi}\left(c(t), g_{ki}(t)\right) \\ = \quad & \mathcal{Q}(t)\, W(t)\, A_{ki}(t)\, Z_{ki}(t)^{1-\eta} \quad c(t)^{1-\eta\epsilon}\, e^{-\eta\, g_{ki}(t)} \left[e^{g_{ik}(t)} \frac{\eta}{\eta-1} - 1 \right] \end{aligned}$$

where we define the profits $\tilde{\Pi}$ depending only on price gaps and aggregate consumption. Thus we can write the expected discounted profits (not taking into account menu or observation costs) for firm k with initial price gap vector $p_k$as:

$$\mathbb{E}\left[\int_0^\infty \mathcal{Q}(t) \left(\sum_{i=1}^n W(t)\ Z_{ki}(t)^{1-\eta} A_{ki}(t)\ \tilde{\Pi}\left[c(t), g_{ki}(t)\right] \right) dt \mid g_k(0) = g_k \right] \text{ w/price gap evolving as}$$

$$g_{ki}(t) = g_{ki} - \log \frac{Z_{ki}(t)}{Z_{ik}(0)} - \log \frac{W(t)}{W(0)} + \sum_{\tau_j < t} \Delta g_{ki}(\tau_j) \quad \text{for each product } i = 1, ..., n$$

Then

$$\mathcal{V}(\boldsymbol{p_k}, \mathbf{c}; p_k) \equiv \mathbb{E}\left[\int_0^\infty e^{-(r+\pi)\, t} \left(\sum_{i=1}^n W(t)\, c(t)^{1-\epsilon\eta}\, \hat{\Pi}\left(g_{ik}(t)\right) \right) dt \mid g(0) = g \right] \tag{44}$$

where we define the normalized profits as $\hat{\Pi}(g) \equiv e^{-\eta\, g} \left[e^{g} \frac{\eta}{\eta-1} - 1 \right]$

Conduction an expansion in equation (44) around zero price gaps and zero aggregate shock (i.e. steady state consumption), we obtain equation (42).

Price gaps, state space, and aggregate shocks. First we consider the case of zero inflation $\pi = 0$. The combination of the different assumptions give that: i) idiosyncratic shocks to cost, and hence price gaps, are drift-less random walks, ii) steady state inflation is zero ($\pi = 0$), and ii) strategic complementarities of aggregate consumption don't (first order) affect optimal decision rules. In turn, i)-iii) imply that, as stated above, both in state dependent and time dependent models, when prices are adjusted, the price gap is closed. Also, as a corollary, in state dependent models the state for problem of the firm is given by the n-dimensional vector of the price gaps, and the inaction set have relatively simple form. For instance, for $n = 1$ product, the inaction set is a interval, and for $n > 1$ when the elasticities are the same $\varrho = \eta$, and uncorrelated shocks across products ($\bar{\sigma} = 0$), it is an hypersphere.

Furthermore, after a once and for all shock aggregate nominal shock starting from a steady state, we have that iv) equilibrium nominal wages once and for, and v) equilibrium nominal interest rates are constant, Thus, i)-v) imply that an for impulse response function we can assume that the decision rules of the firms stay the same before and after the aggregate shock. In particular, to compute the price level, the only effect is to instantaneously and simultaneously for all firms and products, price gaps are reduced by the same percentage, and subsequently price changes are given by the same decision rules as in steady state.

A.2 GE version of impulse response function

We can use the general equilibrium model of Appendix A where equation (16) is produced by a path of money $\log M(t) = \log \bar{M}(t) + \delta$ for all $t \geq 0$, where $\bar{M}(t) = \bar{M}e^{\pi t}$ is the pre-shock expected path of money, with level $\bar{M}$ right before the shock at time $t = 0$. In this ax Using that the labor market is frictionless, so households are in their labor supply, one can obtain that output can we written as function or real wages, which instead can we written as:

$$\log \frac{c(t)}{\bar{c}} = \frac{1}{\epsilon} \left(\delta - \log \frac{P(t)}{\bar{P}(t)} \right) \tag{45}$$

where $\bar{c}$ is the constant flexible price equilibrium output and where $P(t)$ is the ideal price index at time $t \geq 0$ and $\bar{P}(t)$ is the path of the price level in the steady state before the shock, with $\bar{P}(t) = e^{\pi t}\bar{P}$ for all $t \geq 0$.

Using that the labor market is frictionless, so households are in their labor supply, one can obtain that output can we written as function or real wages, which instead can we written as:

$$\log \frac{c(t)}{\bar{c}} = \frac{1}{\epsilon} \left(\delta - \log \frac{P(t)}{\bar{P}(t)} \right) \tag{46}$$

where $\bar{c}$ is the constant flexible price equilibrium output and where $P(t)$ is the ideal price index at time $t \geq 0$ and $\bar{P}(t)$ is the path of the price level in the steady state before the shock, with $\bar{P}(t) = e^{\pi t}\bar{P}$ for all $t \geq 0$.

For tractability reasons, i.e. to reduce the size of the firm's state space, in the general equilibrium version we assume that idiosyncratic firms cost shocks follow a random walk. In the frictionless case this will imply that relative shares of goods will not be stationary. To avoid this we assume that the preference shocks satisfy $A_{ki}(t) = Z_{ki}(t)^{\eta-1}$ so that, the share of expenditure on different goods are constant in the frictionless case.[42] In this case we have that aggregate consumption $c(t)$ is given by equation (24).

B Proofs

Proof. (of Lemma 1) Here we use that the invariant density of Brownian motion is continuous on the (closure) of the inaction set, i.e. on $\{g \in \mathbb{R}^n : b(g) \le 0\}$. This continuity follows as long as $\sigma > 0$. Our proof strategy is to fix a boundary given by the function $b(\cdot;\Delta)$, and considers a discrete time, discrete state representation, with a time period of length Δ. We write $b(g;\Delta)$ so that in the discrete time version we let $b(g;\Delta) > 0$ for any point that is outside the inaction set and $b(g;\Delta) \le 0$ for those inside. In particular we develop the discrete time version of the Kolmogorov forward equation for the density, i.e. a difference equation in the probabilities evaluated finitely many values of g, denoted by $f(g;\Delta)$. We establish that for a value of g for which $b(g) = 0$, then $\lim_{\Delta \downarrow 0} f(g;\Delta) = 0$.

We will consider a discrete time discrete state space representation of the vector of price gaps. Time periods are of length Δ and thus given by $t = s\Delta$ for non-negative integers $s = 1, 2, \ldots$. The state space is given by an equally space grid with the same step size $\sigma\sqrt{\Delta}$ in each of the n dimensions. Thus in each dimension the price gap takes the values $j\,\sigma\,\sqrt{\Delta}$ for the integers $j = 0, \pm 1, \pm 2, \ldots$. To describe the law of motion of $\{g(t)\}$ we will use $n+2$ random variables in each period. These random variables are i.i.d. trough time, and independent of each other. The first two random variable $q(t)$ is used to model the importance of the common component relative to the idiosyncratic component of the price gap. The variable $\bar{w}(t)$ is used to model the innovations on the common component of each price gap. The remaining n random variables $\{\tilde{w}_{it}\}_{i=1}^{n}$ are used to model the innovations on the idiosyncratic component of each of the n price gaps. The distribution of the random variables are:

$$q(t) = \begin{cases} 0 & \text{with probability } 1-\varrho \\ 1 & \text{with probability } \varrho \end{cases}$$

The random variable $\bar{w}(t)$ is distributed as:

$$\bar{w}(t) = \begin{cases} +1 & \text{with probability } \frac{1}{2}\left[1 - \frac{\pi\sqrt{\Delta}}{\sigma}\right] \\ -1 & \text{with probability } \frac{1}{2}\left[1 + \frac{\pi\sqrt{\Delta}}{\sigma}\right] \end{cases}$$

[42] This is a common assumption, as in Midrigan (2011), Woodford (2008) and others

and for each $i = 1, \ldots, n$ we have that each of the random variables $w_i(t)$ are distributed as:

$$\tilde{w}_i(t) = \begin{cases} +1 & \text{with probability } \frac{1}{2}\left[1 - \frac{\pi\sqrt{\Delta}}{\sigma}\right] \\ -1 & \text{with probability } \frac{1}{2}\left[1 + \frac{\pi\sqrt{\Delta}}{\sigma}\right] \end{cases}$$

Thus we have that each price gap $i = 1, \ldots, n$ has changes given by:

$$g_i(t+\Delta) - g_i(t) = \sqrt{\Delta}\,\sigma\ \left[\, q(t+\Delta)\, \bar{w}(t+\Delta)\ +\ (1 - q(t+\Delta))\, \tilde{w}_i(t+\Delta)\, \right]$$

In words: with probability ϱ the price gaps of all products $i = 1, 2, \ldots, n$ move either all up all down together. With probability $1 - \varrho$ the ups and down movement are independent across products. In the case of an up or a down movements the steps are always of the same size, but the probabilities of the up and down are adjusted away from 1/2 to take the negative drift into account. With these definitions we have:

$$\begin{aligned} \mathbb{E}\left[g_i(t+\Delta) - g_i(t)\right] &= -\pi\Delta \text{ for all } i = 1, 2, \ldots, n, \\ \mathbb{E}\left[(g_i(t+\Delta) - g_i(t))^2\right] &= \sigma^2\Delta \text{ for all } i = 1, 2, \ldots, n, \\ \mathbb{E}\left[(g_i(t+\Delta) - g_i(t))\,(g_j(t+\Delta) - g_j(t))\right] &= \varrho\,\sigma^2\Delta \text{ for all } i \neq j = 1, 2, \ldots, n. \end{aligned}$$

Take any $g' \neq g^*$, so that g' is not the optimal return point. Then the mass in g comes from adjacent points in the state space which belong to the inaction set, i.e.:

$$f\left(g'; \Delta\right) = \sum_{\left\{g\,:\, g'_i = g_i \pm \sqrt{\Delta}\sigma\,,\, i=1,\ldots,n\right\}} f\left(g; \Delta\right)\ 1_{\{b(g;\Delta)\leq 0\}}\ \Pr\left\{g'_1 = g_1 \pm \sqrt{\Delta}\sigma, \cdots, g'_n = g_n \pm \sqrt{\Delta}\sigma\right\} \tag{47}$$

The indicator makes sure that only mass that comes from points within the range of inaction can transit from g to g'. Note that at most mass could come from 2^n different points, since in each dimension g_i could have either increase or decrease.

The same steps than for the one dimensional case apply to the general $n > 1$ dimensional case. For simplicity we concentrate first on the case of independence shocks across the products. In this case, we will take a value of g' for which $b(g'; \Delta) = 0$. For $\Delta > 0$ but small enough, there will be some state g for which $g'_i = g_i \pm \sqrt{\Delta}\sigma$ and for which $b(g) > 0$. In words, the point g' of the state space has fewer than 2^n adjacent points that belong to the range of inaction that can move to g' in exactly one time period of length Δ. We thus have that if

$b(g';\Delta) = 0$ then $\#\left\{g \,:\, g'_i \,=\, g_i \pm \sqrt{\Delta}\sigma\right\} < 2^n$.

$$f(g';\Delta) \leq (1-\varrho)\left(\frac{1}{2}\left[1+\frac{\pi\sqrt{\Delta}}{\sigma}\right]\right)^n \sum_{\left\{g\,:\,g'_i\,=\,g_i\pm\sqrt{\Delta}\sigma\,,i=1,\ldots,n\right\}} f\left(g;\Delta\right)\ 1_{\{b(g;\Delta)\leq 0\}}$$
$$+\,\varrho\left[1+\frac{\pi\sqrt{\Delta}}{\sigma}\right]\left[f\left(g'-\sqrt{\Delta}\sigma;\Delta\right)+f\left(g'+\sqrt{\Delta}\sigma;\Delta\right)\right]$$

There are two reasons for the inequality. The first is that we use for all changes the one with higher probability, the down step. The second is that we disregard the possibility that $b(g) > 0$ when either all the price gaps move up or down together (so there is no indicator in the second term of the right hand side). For small enough Δ we have:

$$\sum_{\left\{g\,:\,g'_i\,=\,g_i\pm\sqrt{\Delta}\sigma\,,\,i=1,\ldots,n\right\}} 1_{\{b(g;\Delta)\leq 0\}} \leq 2^n - 1$$

so that there is at least one state which uncontrolled will move to g' but that it doesn't belong to the inaction set. Thus taking limits:

$$\lim_{\Delta\downarrow 0} \sum_{\left\{g\,:\,g'_i\,=\,g_i\pm\sqrt{\Delta}\sigma\,,i=1,\ldots,n\right\}} 1_{\{b(g;\Delta)\leq 0\}} \leq 2^n - 1$$

Moreover for those g for which $g'_i = g_i \pm \sqrt{\Delta}\sigma$ for all $i = 1,\ldots,n$ we have, by the assumed continuity of the density in the closure of the range of inaction, that:

$$\lim_{\Delta\downarrow 0} f\left(g'_1 \pm \sqrt{\Delta}\sigma,\ldots,g'_n \pm \sqrt{\Delta}\sigma;\Delta\right) = \lim_{\Delta\downarrow 0} f(g';\Delta) = f(g')$$

Hence we have that taking limits on equation (47) for values of g' for which $b(g';\Delta) = 0$:

$$\begin{aligned}
&f(g') = \lim_{\Delta\downarrow 0} f(g';\Delta)\\
&\leq (1-\varrho)\lim_{\Delta\downarrow 0}(1-\varrho)\left(\frac{1}{2}\left[1+\frac{\pi\sqrt{\Delta}}{\sigma}\right]\right)^n \sum_{\left\{g\,:\,g'_i\,=\,g_i\pm\sqrt{\Delta}\sigma\right\}} f\left(g;\Delta\right)\ 1_{\{b(g;\Delta)\leq 0\}}\\
&+\varrho\frac{1}{2}\lim_{\Delta\downarrow 0}\left[f\left(g'-\sqrt{\Delta}\sigma;\Delta\right)+f\left(g'+\sqrt{\Delta}\sigma;\Delta\right)\right]\\
&= \varrho\, f(g') + (1-\varrho)\, f(g')\left(\frac{1}{2}\right)^n(2^n-1)\\
&= f(g')\left[1-(1-\varrho)\left(\frac{1}{2}\right)^n\right]
\end{aligned}$$

or $f(g') \leq f(g')\left[1-(1-\varrho)\left(\frac{1}{2}\right)^n\right]$ which again requires that $f(g')=0$. $\square$

Proof. (of Lemma 2) To establish the desired result in the general multi-product case we first develop an expression for and upper bound of $I(\delta)$, denoted by $\bar{I}(\delta)$. We will show that $I(0)=\bar{I}(0)$, that $I(\delta) \leq \bar{I}(\delta)$ and that $\bar{I}'(0)=0$, which implies the desired result $I'(0)=0$. Instead he upper bound function $\bar{I}(\delta)$ is given by:

$$\bar{I}(\delta) = n\int_{-\infty}^{\infty}\cdots\left[\int_{-\infty}^{\infty}\left[\int_{\underline{g}(g_1-\delta,g_2-\delta,\ldots,g_{n-1}-\delta)}^{\underline{g}(g_1-\delta,g_2-\delta,\ldots,g_{n-1}-\delta)+\delta} f(g_1,g_2,\ldots,g_n)dg_n\right]dg_{n-1}\cdots\right]dg_1$$

where the set $\mathbb{G}_{n-1}$ and the function $\underline{g}:\mathbb{G}_{n-1}\to\mathbb{R}$ are defined as:

$$\mathbb{G}_{n-1} \equiv \left\{g_1,g_2,\ldots,g_{n-1}\in\mathbb{R}^{n-1} : \exists\, g_n\in\mathbb{R} \text{ such that } (g_1,g_2,g_3,...,g_n)\in\mathcal{I}\right\}$$
$$\underline{g}(g_1,g_2,\ldots,g_{n-1}) \equiv \min_x\left\{x:(g_1,g_2,\ldots,g_{n-1},x)\in\mathcal{I}\right\} \text{ for } (g_1,g_2,\ldots,g_{n-1})\in\mathbb{G}_{n-1}$$

$\bar{I}(\delta)$ as n times the number of firms that adjust its price on impact because one price has gotten below the lower sS bound (for simplicity we have taken this to be the n price gap, but given exchangeability it does not matter which one it is) . The reason why $\bar{I}(\delta)$ in an upper bound of $I(\delta)$ is that in $\bar{I}$ there is some double counting. The double counting comes from the fact that for some values of g there may have been lower boundaries corresponding to more than one price gap that are crossed after the shock δ. This establishes that $I(\delta)\leq\bar{I}(\delta)$. That $\bar{I}(0)=0$ it follows directly from its definition, since the last intergral is performed in a degenerate interval if $\delta=0$. Finally we have:

$$\bar{I}'(\delta) = n\int_{-\infty}^{\infty}\cdots\left[\int_{-\infty}^{\infty}\left[\frac{\partial}{\partial\delta}\int_{\underline{g}(g_1-\delta,g_2-\delta,\ldots,g_{n-1}-\delta)}^{\underline{g}(g_1-\delta,g_2-\delta,\ldots,g_{n-1}-\delta)+\delta} f(g_1,g_2,\ldots,g_n)dg_n\right]dg_{n-1}\cdots\right]dg_1$$

with

$$\begin{aligned}&\frac{\partial}{\partial\delta}\int_{\underline{g}(g_1-\delta,g_2-\delta,\ldots,g_{n-1}-\delta)}^{\underline{g}(g_1-\delta,g_2-\delta,\ldots,g_{n-1}-\delta)+\delta} f(g_1,g_2,\ldots,g_n)dg_n\\ &= f\left(g_1,g_2,\ldots,\underline{g}\left(g_1-\delta,g_2-\delta,\ldots,g_{n-1}-\delta\right)\right)\end{aligned}$$

which equals zero when evaluated at $\delta=0$, since $f\left(g_1,g_2,\ldots,\underline{g}\left(g_1,g_2,\ldots,g_{n-1}\right)\right)$ is the density at the boundary o f the range of inaction. Thus $\bar{I}'(0)=0$. $\square$

C Random observation cost

In this appendix we describe the set-up for observation cost and signals.

The timeline in Figure 2 describes the structure of the observation cost ψ_o, the associated signal ζ and production cost $z = (z_1, \ldots, z_n)$ that occurs at the time τ_i and at the time of the next observation τ_{i+1}, which is $\min\{\mathrm{T}_i, s_i\}$ periods after the current observation τ_i. In another words, T_i is decided at τ_i.

Immediately after paying the observation cost at time τ_i, the firm learns the current value of $(z_1, \ldots, z_n)$ and receives a signal ζ which is informative about the future realizations of the observation cost ψ_o'. Recall that at time τ_i the firm decides its planned elapsed time until the next observation T_i. Recall that the firm also reviews at an exogenous exponentially distributed time s_i with parameter λ. Thus the time elapsed until next observation will occur at the earliest time between s_i, which is a random variable, or T_i, which as of time τ_i is decided, and thus known, i.e. non-random. Summarizing the time for the $i+1$ observation is given by $\tau_i + \min\{s_i\,, \mathrm{T}_i\}$.

The signal ζ summarizes all the information about the value of the observation cost to be paid $T = \min\{s_i\,, \mathrm{T}_i\}$ periods from now. Mathematically we write $F(\psi_o'\,; T|\zeta)$ to be the CDF of the observation cost ψ_o' to be paid T periods after the current observation, conditional on the signal ζ. The dependence of the distribution F on T allows the distribution of the observation cost ψ_o' to vary with the time elapsed between observations. The functions F and G fully characterize the process for the observation cost, and provide enough flexibility to cover cases discussed in the literature as well as generalizations that we find useful. The expected observation cost is the key input to decide T.

Upon the next observation, when a particular cost ψ_o' is realized, a new signal ζ' is drawn from the CDF $G(\cdot|\psi_o')$. The other key input to decide T is the distribution of $z_j(\tau_{i+1})$ conditional on $z(\tau_i)$, which, given the assumption that the log of $\{z_j\}$ are random walks, we can summarize them as $L(\cdot\,; T\,|\,z)$. These distributions allows to compute the benefit of gathering information, i.e. of choosing a small value of T_i.

In this set-up we can obtain several of the cases analyzed in the literature. For instance, the model of deterministic observation times studied by Caballero (1989) and Reis (2006) is encompassed by our framework if the signal is uninformative about the future observation cost, which is the case if $F(\psi_o', T_0\,|\,\zeta_0) = F(\psi_o', T_1\,|\,\zeta_1)$ for all ψ_o' and all pairs (T_0, ζ_0). In this case, the distribution G is irrelevant because, given that the signal is uninformative, the mechanism to obtain the new signal is irrelevant.

Another case discussed in the literature is one where the firm's observation times are i.i.d., as proposed by Reis (2006). This setup provides a foundation to i.i.d. observation times: the

Figure 2: Time line

τ_i	$\tau_{i+1} = \min\{s_i, \mathrm{T}_i\} + \tau_i$
pay observation cost ψ_o	time between observations $T \equiv \min\{s_i, \mathrm{T}_i\}$
learn production costs $(z_1, \ldots, z_n)$	draw $z_j' \sim L(\cdot\,;\, T \,|\, z)$ for $j = 1, \ldots, n$
learn signal ζ	draw $\psi_o' \sim F(\cdot\,;\, T \,|\, \zeta)$
choose planned T_i	draw $\zeta' \sim G(\cdot \,|\, \psi_o')$

firm has to draw a signal about the future observation cost that is both informative about the next observation cost and independent of all other shocks (including the current value of the observation cost). In this case, the particular form of the distribution G is relevant. Formally, observation times are i.i.d. in our model if and only if $G\left(\zeta \,|\, \tilde{\psi}_o\right) = G\left(\zeta \,|\, \bar{\psi}_o\right)$ for all ζ and all pairs $\tilde{\psi}_o, \bar{\psi}_o$. The distribution F shapes the precision of the signal. Finally, the more general case where $G\left(\zeta \,|\, \tilde{\psi}_o\right) \neq G\left(\zeta \,|\, \bar{\psi}_o\right)$ for at least some ζ and some pairs $\tilde{\psi}_o, \bar{\psi}_o$ allows us to extend our analysis to the case of observation times correlated over time, a case which we find more reasonable than the i.i.d. assumption.

D Time and State dependent firm's problem

In the multi-product Calvo$^+$ model the firms solves:

$$\max_{\{\tau_i, \Delta P_j(\tau_i), j=1,\ldots,n, i=1,2,\ldots,\}} \mathbb{E}\left[\int_0^\infty e^{-rt}\,\Pi\left(P_1(t), ..., P_n(t), Z_1(t), ..., Z_n(t)\,,\, W(t)\,;\, c(t)\right)\, dt \right.$$
$$\left. - \sum_{i=1}^{\infty} e^{-r\tau_i}\,\psi_m\, 1_{\{dU(\tau_i)=1\}}\, W(\tau_i)\right] \tag{48}$$

$P_j(t) = P_j(\tau_i)$ for all $t \in (\tau_i, \tau_{i+1}]$ and $\Delta P_j(\tau_i) = \lim_{\epsilon \downarrow 0} P_j(\tau_i + \epsilon) - P_j(\tau_i)$

where $\{U(t)\}$ is a Poisson process w/intensity λ.

In the problem with observation and menu cost the firm solves

$$\max_{\{\text{T}_i,\, a(\tau_i),\, \Delta P_j(\tau_i),\, j=1,\ldots,n, i=1,2,\ldots\}} \mathbb{E}\left[\int_0^\infty e^{-rt}\,\Pi\left(P_1(t),...,P_n(t),Z_1(t),...,Z_n(t)\,,\,W(t);\,c(t)\right)\;dt \right.$$
$$\left. -\sum_{i=1}^{\infty} e^{-r\tau_i}\,\psi_o\left[1+a(\tau_i)\psi_m\right]W(\tau_i)\right] \tag{49}$$

where $\tau_{i+1} = \text{T}_i + \tau_i$ where T_i is the time between observations,

T_i and $a(\tau_i) \in \{0,1\}$ only depend on information gathered at $\tau_0, \tau_1, \ldots, \tau_i$,

$P_j(t) = P_j(s)$ for all $s, t \in (\tau_i, \tau_{i+1})$ and

$$\Delta P_j(\tau_i) = \begin{cases} \lim_{\epsilon \downarrow 0} P_i(\tau_i+\epsilon) - P_i(\tau_i) & \text{if } a(\tau_i)=1 \\ 0 & \text{if } a(\tau_i)=0 \end{cases}$$

We simplify the problem assuming that both observation and menu costs are non-random.

E Tradeoff of sticky price models around zero inflation.

In all the micro-founded models we consider the nominal price upon adjustment is reset at the optimal price maximizing level, i.e. the price gap is "closed", so that the size of the price adjustment is equal to (minus) the price gap, or $\Delta g_i = -\Delta \log P_i$. In this broad class of models we have that the following property holds for any decision rules and $n \geq 1$:

$$N(\Delta p_i;0)\;Var(\Delta p_i;0) = \sigma^2\;,\;\text{and}\;\frac{\partial}{\partial \pi}\left[N(\Delta p_i;\pi)\;Var(\Delta p_i;\pi)\right]\Big|_{\pi=0} = 0 \tag{50}$$

which states that the total number of price adjustments per period denoted by $N(\Delta p_i;\pi)$ times the variance of the size of price changes, $Var(\Delta p_i;\pi)$ equals the variance of the innovations to the price gaps σ^2. This equation, which holds for any policy for inflation around zero, which end up closing the price gap upon adjustments (even non-optimal policies) highlights the key tradeoff of a sticky price problem, that between the frequency of costly adjustments, or information gathering, versus the mean deviation of nominal prices from their optimal level.

F Sensitivity of impact effect to inflation

In this section we derive the case of a state dependent model with one good $n = 1$. We explicitly write the barriers and the optimal return point as function of inflation, for a given

$\sigma > 0$. At zero inflation we have: $\bar{g}(0) = -\underline{g}(0) > -$ and $g^*(0) = 0$. We will like to obtain an expansion of $I(\delta, \pi)$. We have

$$\begin{aligned} I(\delta; \pi) &= f(\underline{g}(\pi); \pi)\, \delta + \frac{1}{2} f'(\underline{g}(\pi); \pi)\, \delta^2 + \frac{1}{6} f''(\underline{g}(\pi); \pi)\, \delta^3 + o(\delta^3) \\ &= \frac{1}{2} f'(\underline{g}(\pi); \pi)\, \delta^2 + \frac{1}{6} f''(\underline{g}(\pi); \pi)\, \delta^3 + o(\delta^3) \end{aligned}$$

where f' and f'' denote the derivatives of the density with respect to g, and where f'_π and f''_π denote the (cross) derivatives with respect to π. Note that at $\pi = 0$, $f(\cdot, 0)$ is linear so that $f''(g, 0) = 0$. We thus have:

$$\begin{aligned} I(\delta; \pi) &= I(\delta; 0) + \frac{1}{2} \frac{\partial}{\partial \pi} f'(\underline{g}(0); 0)\, \pi\, \delta^2 + o\left(||(\delta, \pi)||^3\right) \\ &= \frac{1}{2} f'(\underline{g}(0); 0)\, \delta^2 + \frac{1}{2} \frac{\partial}{\partial \pi} f'(\underline{g}(0); 0)\, \pi\, \delta^2 + o\left(||(\delta, \pi)||^3\right) \\ &= \frac{1}{2} \frac{1}{\underline{g}(0)^2}\, \delta^2 + \frac{1}{2} \frac{\partial}{\partial \pi} f'(\underline{g}(0); 0)\, \pi\, \delta^2 + o\left(||(\delta, \pi)||^3\right) \end{aligned}$$

To be more precise, the density depends on inflation directly, in the dependence of the Kolmogorov forward equation on π:

$$0 = f'(g; \pi, \underline{g}, \bar{g}, g^*)\pi + f''(g; , \underline{g}, \bar{g}, g^*) \frac{\sigma^2}{2} \text{ for all } g \in [\underline{g}, \bar{g}], g \neq g^* \tag{51}$$

as well as, indirectly on $\underline{g}(\pi), \bar{g}(\pi)$ and $g^*(\pi)$ which depend on π. We have:

$$\begin{aligned} \frac{\partial}{\partial \pi} f'(\underline{g}(0); 0, \underline{g}(0), \bar{g}(0), g^*(0)) &= f''(\underline{g}(0); 0)\, \underline{g}'(0) + f'_\pi(\underline{g}(0); 0) \\ &+ f'_{\underline{g}}(\underline{g}(0); 0, \underline{g}(0), \bar{g}(0), g^*(0))\, \underline{g}'(0) \\ &+ f'_{\bar{g}}(\underline{g}(0); 0, \underline{g}(0), \bar{g}(0), g^*(0))\, \bar{g}'(0) \\ &+ f'_{g^*}(\underline{g}(0); 0, \underline{g}(0), \bar{g}(0), g^*(0))\, g^{*\prime}(0) \end{aligned}$$

where we use that at $\pi = 0$ the density is linear in g, i.e.

$$f(g, 0) = \frac{g - \underline{g}(0)}{\underline{g}(0)^2} \text{ for } g \in [\underline{g}(0), 0) \tag{52}$$

and hence its derivative f' does not depend on g. Thus we have:

$$I(\delta; \pi) = \frac{1}{2} \frac{1}{\underline{g}(0)^2}\, \delta^2 + \frac{1}{2} f'_\pi(\underline{g}(0); 0)\, \pi\, \delta^2 + o\left(||(\delta, \pi)||^3\right)$$

We first study the effect of inflation on f' keeping the thresholds $\underline{g}, \bar{g}$ and the optimal return point g^* fixed as we change inflation. Thus, we only study the direct effect of inflation on f'. To do this, we first determine f and its derivatives. The density f solves the Kolmogorov forward equation at all $g \neq g^* = 0$. Furthermore we use that the density is zero at the exit points. Thus there must be two constants $\underline{B}$ and $\bar{B}$:

$$f(g,a) = \begin{cases} \underline{B}(a)\left(e^{ag} - e^{a\underline{g}}\right) & \text{if } g \in [\underline{g}, g^*] \\ \bar{B}(a)\left(e^{ag} - e^{a\bar{g}}\right) & \text{if } g \in [g^*, \bar{g}] \end{cases} \tag{53}$$

where we use a to denote the non-zero root of the characteristic equation for the solution of the KF equation:

$$a = -2\pi/\sigma^2 \,. \tag{54}$$

We describe the two equations for the constants $\underline{B}(a)$ and $\bar{B}(a)$. Continuity of the density $f(\cdot, a)$ at $g^* = 0$ gives

$$\underline{B}(a)\left(e^{ag^*} - e^{a\underline{g}}\right) = \bar{B}(a)\left(e^{ag^*} - e^{a\bar{g}}\right)$$

and integrates to one:

$$1 = \int_{\underline{g}}^{\bar{g}} f(g,a)dg = \int_{\underline{g}}^{g^*} \underline{B}(a)\left(e^{ag} - e^{a\underline{g}}\right)dg + \int_{g^*}^{\bar{g}} \bar{B}(a)\left(e^{ag} - e^{a\bar{g}}\right)dg$$

We are interested in:

$$f'(\underline{g}(0);0) = \lim_{a\to 0} \underline{B}(a)\, a\, e^{a\underline{g}(0)} = \frac{1}{\underline{g}(0)^2} \text{ and} \tag{55}$$

$$f'_\pi(\underline{g}(0);0) = -\frac{2}{\sigma^2} \lim_{a\to 0} \frac{\partial}{\partial a} \underline{B}(a)\, a\, e^{a\underline{g}(0)}\Big|_{a=0} \tag{56}$$

Solving for $\underline{B}$ we have:

$$\begin{aligned} 1 &= \underline{B}(a)\left(\frac{e^{ag^*} - e^{a\underline{g}}}{a} - e^{a\underline{g}}(g^* - \underline{g})\right) + \bar{B}(a)\left(\frac{e^{a\bar{g}} - e^{ag^*}}{a} - e^{a\bar{g}}(\bar{g} - g^*)\right) \\ &= \underline{B}(a)\left(\frac{e^{ag^*} - e^{a\underline{g}}}{a} - e^{a\underline{g}}(g^* - \underline{g})\right) + \underline{B}(a)\frac{\left(e^{ag^*} - e^{a\underline{g}}\right)}{\left(e^{ag^*} - e^{a\bar{g}}\right)}\left(\frac{e^{a\bar{g}} - e^{ag^*}}{a} - e^{a\bar{g}}(\bar{g} - g^*)\right) \end{aligned}$$

where we use that the density integrates to one and that it is continuous at zero. Then

$$\begin{aligned}
1 &= \underline{B}(a)\left\{\frac{e^{ag^*}-e^{a\underline{g}}}{a} - e^{a\underline{g}}(g^*-\underline{g}) + \frac{e^{ag^*}-e^{a\underline{g}}}{e^{ag^*}-e^{a\bar{g}}}\left(\frac{e^{a\bar{g}}-e^{ag^*}}{a} - e^{a\bar{g}}(\bar{g}-g^*)\right)\right\}\\
&= -\underline{B}(a)\left\{e^{a\underline{g}}(g^*-\underline{g}) + \frac{e^{ag^*}-e^{a\underline{g}}}{e^{ag^*}-e^{a\bar{g}}}\, e^{a\bar{g}}(\bar{g}-g^*)\right\}\\
&= -\underline{B}(a)\left\{\frac{e^{ag^*}-e^{a\bar{g}}}{e^{ag^*}-e^{a\bar{g}}} e^{a\underline{g}}(g^*-\underline{g}) + \frac{e^{ag^*}-e^{a\underline{g}}}{e^{ag^*}-e^{a\bar{g}}}\, e^{a\bar{g}}(\bar{g}-g^*)\right\}\\
&= -\frac{\underline{B}(a)}{e^{ag^*}-e^{a\bar{g}}}\left\{\left(e^{ag^*}-e^{a\bar{g}}\right) e^{a\underline{g}}(g^*-\underline{g}) + \left(e^{ag^*}-e^{a\underline{g}}\right) e^{a\bar{g}}(\bar{g}-g^*)\right\}
\end{aligned}$$

using that $\underline{g} = -\bar{g}$ and $g^* = 0$ at $\pi = 0$:

$$\begin{aligned}
1 &= -\frac{\underline{B}(a)}{e^{ag^*}-e^{a\bar{g}}}(\bar{g}-g^*)\left\{\left(e^{ag^*}-e^{a\bar{g}}\right) e^{a\underline{g}} + \left(e^{ag^*}-e^{a\underline{g}}\right) e^{a\bar{g}}\right\}\\
&= -\underline{B}(a)\frac{(\bar{g}-g^*)}{e^{ag^*}-e^{a\bar{g}}}\left\{e^{ag^*}e^{a\underline{g}} - 1 + e^{ag^*}e^{a\bar{g}} - 1\right\}\\
&= -\underline{B}(a)\frac{\bar{g}}{1-e^{a\bar{g}}}\left\{e^{-a\bar{g}} - 1 + e^{a\bar{g}} - 1\right\}
\end{aligned}$$

or

$$\underline{B}(a) = -\frac{(1-e^{a\bar{g}})}{\bar{g}\left(e^{-a\bar{g}} - 1 + e^{a\bar{g}} - 1\right)} < 0 \text{ since } a < 0$$

We thus have:

$$\begin{aligned}
\bar{g}^2 f'(\underline{g}, a) = C(\alpha) \equiv \bar{g}^2\, \underline{B}(a)\, a\, e^{-a\bar{g}} &= -\frac{\alpha\left(e^{-\alpha}-1\right)}{\left(e^{-\alpha}-1+e^{\alpha}-1\right)}\\
&= -\left(\frac{1}{2}\right)\frac{-\alpha^2 + \alpha^3/2 - \alpha^4/3! + \cdots}{\alpha^2/2 + \alpha^4/4! + \alpha^6/6! + \cdots}\\
&= -\left(\frac{1}{2}\right)\frac{-1 + \alpha/2 - \alpha^2/3! + \cdots}{1/2 + \alpha^2/4! + \alpha^4/6! + \cdots}\\
&\text{where } \alpha \equiv a\,\bar{g} < 0
\end{aligned}$$

Note that direct computation gives

$$C(0) = 1 \text{ and } C'(0) = -\frac{1}{2}$$

Thus:

$$f'(\underline{g},\pi) = \frac{1}{\bar{g}^2}\,C\left(-\frac{2\pi}{\sigma^2}\bar{g}\right) \to \frac{1}{\bar{g}^2} \text{ and}$$
$$f'_\pi(\underline{g},0) = -\frac{2\,\bar{g}}{\sigma^2\bar{g}^2}C'\left(0\right) = \frac{1}{\sigma^2\bar{g}}$$

G Accuracy of Proposition 2 in models with both menu and observation costs

We numerically evaluate the accuracy of the approximation in Proposition 2 in models that feature the simultaneous presence of a menu cost $\psi_m > 0$ as well as an observation cost $\psi_o > 0$. The main parameter to be specified for this analysis is the ratio between the two fixed costs: $\alpha \equiv \psi_m/\psi_o$. Notice that for the special case of the observation cost only ($\psi_m = 0$ so that $\alpha = 0$) as well as the special case of the menu cost only ($\psi_o = 0$ so that $\alpha \to \infty$), we have supplied an analytic proof of the proposition in Alvarez, Le Bihan, and Lippi (2016) for $\psi_m > 0, \psi_o = 0$ and in Alvarez, Lippi, and Paciello (2016) for $\psi_o > 0, \psi_m = 0$.

To analyze the problem with $0 < \alpha < \infty$ we use the decision rules derived in Alvarez, Lippi, and Paciello (2011) and numerically compute the invariant distribution of firms in a steady state. This is a joint density defined over the time until the next review and the value of each firm price gap. We then develop the impulse response analysis by shocking the steady state of the economy and computing the are under the impulse response. There is essentially one parameter in this analysis, α, since the two fixed costs enter the problem as a ratio and the policy functions are homogenous, so that the results only depend on the ratios of particular moments (e.g. frequency of adjustment versus frequency of observation; see Alvarez, Lippi, and Paciello (2011) for details). In practice, we normalize the value of $N = 1$ in all the models we consider and vary α so that the models will generate different steady state levels of kurtosis.

Figure 3 summarizes the results of our numerical analysis. The vertical axis plots the ratio of the area under the output impulse, numerically computed, and the approximation of the same object given by the ratio of the steady state moments $Kurt(\Delta p_i)/N(\Delta p_i)$ as suggested in Proposition 2. We consider a set of models where $0 < \alpha < 5$. The model uses a weekly time period and a cross section of 100 thousand firms, and a monetary shock equal to 1 per cent i.e. $\delta = 0.01$. It appears that the numerical accuracy of the proposition is within $\pm 5\%$ of the actual cumulative effect and, more importantly, that the accuracy does not display a systematic variation with respect to α.

Figure 3: Ratio between actual and approximate cumulative output effect