DISCUSSION PAPER SERIES

No. 9228

MONETARY SHOCKS IN A MODEL WITH INATTENTIVE PRODUCERS

Fernando E Alvarez, Francesco Lippi and Luigi Paciello

INTERNATIONAL MACROECONOMICS

Centre for Economic Policy Research

www.cepr.org

Available online at: www.cepr.org/pubs/dps/DP9228.asp

ISSN 0265-8003

MONETARY SHOCKS IN A MODEL WITH INATTENTIVE PRODUCERS

Fernando E Alvarez, University of Chicago and NBER
Francesco Lippi, Università di Sassari, EIEF and CEPR
Luigi Paciello, EIEF

Discussion Paper No. 9228
November 2012

Centre for Economic Policy Research
77 Bastwick Street, London EC1V 3PZ, UK
Tel: (44 20) 7183 8801, Fax: (44 20) 7183 8820
Email: cepr@cepr.org, Website: www.cepr.org

CEPR Discussion Paper No. 9228

November 2012

ABSTRACT

Monetary Shocks in a Model with Inattentive Producers*

We study a model in which prices respond slowly to shocks because firms must pay a fixed cost to observe the determinants of the profit maximizing price, as pioneered by Caballero (1989) and Reis (2006). We extend their analysis to the case of random transitory variation in the firm's observation cost and characterize the mapping from the distribution of observation cost to the distribution of the times between consecutive re- views/price adjustments of a firm. We aggregate a continuum of firms and characterize analytically the cross-sectional distribution of the duration of reviews/prices. We establish the dependence of the real effect of a monetary shock on the distribution of price durations and hence on the distribution of observation costs and discuss applications.

JEL Classification: e5
Keywords: impulse responses., inattentiveness, monetary shocks and observation costs

Fernando E Alvarez
Department of Economics
University of Chicago
1126 East 59th Street
Social Science Building, Room 442
Chicago IL, 60637
USA

Email: f-alvarez1@uchicago.edu

For further Discussion Papers by this author see:
www.cepr.org/pubs/new-dps/dplist.asp?authorid=137370

Francesco Lippi
Universita Degli Studi di Sassari
Department of Economics
Via Torre Tonda, 34
07100 Sassari
ITALY

Email: f.lippi@uniss.it

For further Discussion Papers by this author see:
www.cepr.org/pubs/new-dps/dplist.asp?authorid=141851

Luigi Paciello
Einaudi Institute for Economics and Finance
Via Sallustiana, 62
00187 Rome
ITALY

Email: luigi.paciello@eief.it

For further Discussion Papers by this author see:
www.cepr.org/pubs/new-dps/dplist.asp?authorid=169105

*First draft, January 2012. We thank Carlos Carvalho, Xavier Gabaix, Christian Hellwig, Bartosz Mackowiak, Galo Nuno, Mirko Wiederholt, as well as seminar participants at the ECB, NBER Summer Institute 2012 and Tor Vergata for their comments.

Submitted 14 November 2012

1 Introduction

The pioneering contributions of Caballero (1989) and Reis (2006) laid the foundations of the models where firms adjust prices infrequently because of an underlying cost of gathering information. In these models the firms must pay a fixed cost to observe or "review" the determinants of the profit maximizing price, and change prices in response to shocks only when they gather the relevant information. Both Caballero and Reis provide a characterization of the firm's decision rule assuming the observation cost is constant so that, given the problem setup, the time elapsed between the reviews is constant. Reis (2006) also provides a setup for studying the aggregation of a cross section of firms, each of which follows an exogenous stochastic distribution of review times.

The first contribution of this paper is to provide a foundation to the stochastic review times by studying a firm's decision about when to review and change prices in the face of random transitory variation in observation costs. This is important because, compared to the constant review times, the stochastic reviews have been shown to matter for the aggregate effect of monetary policy e.g. in Mankiw and Reis (2002) and more generally in Carvalho and Schwartzman (2012). We derive analytically the implications of a given distribution of observation costs for: (i) the distribution of the times elapsed between consecutive reviews for an individual firm, and (ii) an aggregation result that gives the invariant distribution of the times until the next review for a cross section of firms, what Reis labels "distribution of inattentiveness". We also establish (iii) a simple mapping that links the cross sectional distribution of price changes to the distribution of the review times. These results are used to characterize (iv) how the cumulative output response to a monetary shock depends on the distribution of the observations cost. Our analytical results highlight what assumptions shape the real effect of monetary shocks in costly observation models.[1] The linkages that are unveiled between e.g. the size distribution of price changes and the distribution of the observation costs provide a metric that could be used to test the model assumptions empirically.

A second contribution of the paper is to clarify two related theoretical results in Reis (2006) concerning the aggregation of individual firms decision. Assuming iid review times at the firm level Reis claims that, for any distribution of the firm level review times, the cross sectional distribution of inattentiveness will be exponential. This result is important because exponentially distributed review times yield large real effects of monetary shocks, as in e.g. Mankiw and Reis (2002) and others. We show that Reis's results are incorrect. The condition to obtain an exponential distribution of review times is restrictive: the firm level

[1] In a companion paper, Alvarez, Lippi, and Paciello (2012), we show that the analytical results provide a good approximation to the ones obtained in more involved quantitative models which cannot be solved analytically.

distribution of review times must be exponential itself.

Next we give an overview of the model setup and the main results. The rest of the paper is organized as follows. In Section 2 we describe the firm's problem and characterize the optimal decision rules. Section 3 studies the cross-sectional distribution of review times implied by the firm's decision rules as well as the distribution of observation costs that can rationalize a given cross-sectional distribution of review times. Section 4 shows how the cross-sectional distribution of review times shapes the response of the economy to a monetary shock. Section 5 revises Reis's (2006) propositions about aggregation. Section 6 discusses an alternative formulation of the model which assumes *permanent* cross section differences in the review times. Section 7 concludes.

Overview and main results

We consider a simple general equilibrium model of price setting where firms choose the time of price reviews facing transitory idiosyncratic variation in observation costs. We focus on transitory (iid) variation of the observation cost mostly for a literature-related reason, namely to generate firm price reviews and adjustments that are iid through time. Throughout the paper we assume that there are no menu costs so that each price review yields a price adjustment.[2] An example of this is the Calvo mechanism, as well as the study of aggregation in Reis (2006), among others. Furthermore to endogenously generate iid times to review the rational inattentive model requires either cost or benefit of observations that are iid across review-adjustments. We derive the mapping from the distribution of observation costs, $w(\theta)$, to the distribution of optimally chosen time elapsed between consecutive reviews, $H(\tau)$. Next, we derive the mapping from the distribution $H(\tau)$ for an individual firm to the invariant distribution of the times r until the next review for a cross section of firms, $Q(r)$, what Reis labels the "distribution of inattentiveness". Furthermore we derive a simple relationship between the distribution of inattentiveness and the cumulative output response to a monetary shock, and between the distribution of inattentiveness and the distribution of price changes. Our analysis is related to the ongoing work by Carvalho and Schwartzman (2012) who study how different distributions of price durations map into different real effects of monetary shocks. While in their analysis this distribution is a primitive ingredient, our paper endogenizes it as a function of the distribution of observation costs. Our analysis can thus be used to identify what variability of the observation costs is needed to support a given variability of price spells and, therefore, a given persistence of of aggregate price level and

[2] Moreover we do not let the firm choose a price plan: as shown in Proposition 1 of Alvarez, Lippi, and Paciello (2011), relatively small menu costs make price-plans not optimal in this environment at moderate inflation rates.

output to a permanent unexpected increase in money supply.

The analysis yields the following results. First, while with a constant observation cost the optimal time to the next review is a strict *concave* function of the cost, under transitory idiosyncratic variation in observation costs the firm's optimal time to the next review is a *convex* function of the cost. The reason is that the variability of observation costs, and therefore the uncertainty about the cost of the next review, introduces an option value of waiting, so that at low values of the cost the firm finds it optimal to wait before observing and drawing a new observation cost. As a consequence, the stochastic observation cost induces the firm to optimally choose a strictly positive duration of time between reviews even when it draws a zero cost for the next observation. Thus, the model supports arbitrarily small time between reviews if and only if there is a positive mass of *negative* observation cost. Moreover, we show that the introducing a small variability of observation costs has a negligible effect on the variability of time elapsed between reviews. This last result is important to map the variability of observation costs into the real output response to monetary shocks, discussed below.

Second, contrary to what is stated in Reis (2006), we show that the distribution of inattentiveness, $Q(r)$, is exponential if and only if the distribution of the firm's times between consecutive reviews, $H(\tau)$, is exponential. In addition, we derive a mapping that identifies what distribution of the observation costs $w(\theta)$ yields a desired distribution $H(\tau)$.

Third we characterize the real effect of a monetary shock. It is shown that the cumulated response of output to a monetary shock (the area under the impulse response of output) is a function of only two moments of the distribution $H(\tau)$: the mean (i.e. average duration) and the coefficient of variation of the time elapsed between reviews. The cumulated output response is increasing in the average duration, as well as in the coefficient of variation of durations. Keeping fixed the average time between reviews, the cumulated output response in a model with exponentially distributed reviews, as in Mankiw and Reis (2002), is twice as big as the one in a model with a constant observation cost yielding a degenerate distribution of review times, as in Caballero (1989), Reis (2006) and Bonomo and Carvalho (2004). The fact that the variability of observation cost has a small effect on the variability of the time between reviews has interesting implications for the type of observation costs that are need to produce large output effects. For instance to obtain effects of monetary shocks such as those obtained with an exponential distribution of inattentiveness, with parameter values chosen to match the frequency and size of price changes from the US CPI data, the model requires a large mass of *negative* observation costs as well as very large variation in observation costs.

In the sticky price literature impulse responses with exactly those shapes can be obtained assuming a *Calvo* mechanism, i.e. an exponential distribution of times between adjustments, or a *Taylor* "staggered" price setting rule, which implies a degenerate distribution of times between adjustments. Interestingly Chari, Kehoe, and McGrattan (2000) have found that a sticky price model with a *Taylor* mechanism predicts a much smaller real effect than with a *Calvo* mechanism. Hence the findings concerning the real effects of monetary shocks in sticky information models are reminiscent of their findings about the sticky price models.

In Alvarez, Lippi, and Paciello (2011) we have studied an analytical price setting problem in the presence of disjoint observation and adjustments costs. In such a setup not all reviews yield a price adjustments and, therefore, produce variability in the times elapsed between reviews.[3] In that model, the time elapsed between price adjustments/reviews depends on the value of the price gap upon an observation and is, therefore, state-dependent. In Alvarez, Lippi, and Paciello (2012) we study the impulse responses for the economy with disjoint constant observation and adjustment costs and find that the size of the real effects of a monetary shock is *even* smaller than in the model of this paper where price reviews and adjustments are iid through time.

In the spirit of the "rational inattention" literature that originated from Sims (2003), Mackowiak and Wiederholt (2009, 2011) have studied a model where firms face a constraint in the amount of information they can process in each period and choose how much attention to devote to aggregate versus idiosyncratic shocks. They have shown that, assuming a small cost of the information processing capacity, firms pay relatively little attention to aggregate shocks so that monetary shocks can have large real effects. We show that the structure of observation costs that produces large real effects in our model requires large variability and negative values of the observation costs. We think that exploring the mapping between the observation costs, or the attention allocation in "rational inattention" models, into directly observable actions (or costs borne) by the firm is an interesting avenue for future research.

2 Optimal price setting with random observation costs

This section presents the price setting problem of a firm subject to random i.i.d. observation costs. The goal is to understand the mapping between the properties of the distribution of observation cost and the firm's optimal decision about when to review the relevant information for its price decision, and adjust the price. This will be used, for instance, to study what distribution of the observation costs implies an exponential distribution of the times until the next review for the cross section of firms, as assumed e.g. in the *inattentiveness* model of Mankiw and Reis (2002).

[3] See Bonomo, Carvalho, and Garcia (2010) and Burstein (2006) for related analysis of the price setting choices with disjoint costs and the possibility of price plans.

We consider a set-up where competitive monopolistic firms are subject to random shocks to their productivity as well as to the cost to observe their own productivity to review the prices. We select a specification of productivity and observation costs to obtain that successive price reviews are i.i.d. through time. We focus on transitory (i.i.d.) variation of θ mostly for a literature-related reason, namely to endogenously generate a setup where price reviews (and adjustment) are i.i.d. through time for a given firm, as assumed for instance in Mankiw and Reis (2002). In particular, in the context of endogenously inattentive firms as modeled by Caballero (1989) and refined by Reis (2006), this means that i.i.d. variation in reviews should stem either from the costs or from the benefits of a review. In fact, variation in production cost or demand would not, in general, produce i.i.d. reviews since typical shocks to either the production technology or the demand function would produce persistent effects on the incentives of a review, so also affecting the next review decision.[4]

We first describe the problem of the firm in detail. The monopolist faces a demand with constant elasticity η and a constant returns to scale production where the (log) of the labor productivity z follows a brownian motion with drift γ and volatility σ. Nominal wages, W_t, as well as the average price of the rest of the goods increase at the constant rate of inflation μ, and the nominal interest rate is constant and equal to $\rho + \mu$.[5] It is also assumed that the product of the firm is replaced ("dies") with a constant probability per unit of time equal to λ, and the new product is born with productivity $z = 1$. We interpret the substitution of a product as triggering an observation corresponding to the a new product that has replaced it. When this happens a new good is born and it is optimally priced. More technically, exogenous substitutions allow for an invariant distribution of firm's productivities and, therefore, a well defined measure of aggregate output. The firm's nominal per period profits are given by

$$C_t(p)\,(p - W_t/z)\ ,$$

where $C_t(p) = A_t\ \left(\frac{p}{W_t}\right)^{-\eta}$ is the downward sloping real demand function, A_t is deterministic and proportional to the level of aggregate demand, and $\Pi_t^*(z)$ is the level of profits at the period profit maximizing price, i.e. [6]

[4]In Alvarez, Lippi, and Paciello (2012) we study a model where the presence of both observation and menu costs produces state-dependent variation in review times. Woodford (2009) assumes that when firms decide to conduct a new review they cannot recall the state of the past observation. With the assumption of no recall of any past information, reviews are iid in a stationary environment.

[5] Section 4 presents a general equilibrium model where this path of nominal wages arises in equilibrium.

[6]The subscript t in $\Pi_t(p, z)$ captures the variation in profits due to the dynamics of A_t and W_t. The real demand function $C_t(p) = A_t \left(\frac{p}{W_t}\right)^{-\eta}$ can be derived from a general equilibrium model where a representative household consumes different varieties, with elasticity of substitution η. With linear disutility in labor, and intertemporal elasticity of substitution in consumption equal to $1/\epsilon$, the intercept is $A_t = c_t^{1-\eta\epsilon}$ where c_t is the aggregate consumption. See Alvarez, Lippi, and Paciello (2012) for more details.

$$p_t^* = \frac{\eta}{\eta-1}\frac{W_t}{z} = \arg\max_p \Pi_t(p,z) \ . \tag{1}$$

Using equation (1), we obtain that profits, scaled by the nominal wages, are given by:

$$\Pi_t(p,z) \ = \ A_t \ z^{\eta-1} \ F\left(p/p^*\right) \qquad \text{and} \qquad \Pi_t^*(z) = A_t \ z^{\eta-1} \ F(1) \tag{2}$$

where, letting $g \equiv \frac{p}{p^*}$, the function $F(\cdot)$ is

$$F(g) \equiv \left(g \, \frac{\eta}{\eta-1}\right)^{-\eta}\left(g \, \frac{\eta}{\eta-1} - 1\right) \quad . \tag{3}$$

We refer to the ratio g of the nominal price relative to the current value of the optimal static price for the monopolist as to the "price gap". A value of 1 indicates that the current sale price is the optimal (static) one. It follows from the definition of g, and from the laws of motion of W and z, that the dynamics of g are given by

$$d\log(g(t)) = (\gamma - \mu) \ dt + \sigma dB(t), \tag{4}$$

where $B(t)$ is a Wiener process. Next, define the function $f(g,s)$ as:

$$f(g,s) \ \equiv \ \mathbb{E}\left[\left(\frac{z_{t+s}}{z_t}\right)^{\eta-1}\frac{F(g_{t+s})}{F(1)}\Big|g_t = g\right] = \eta g^{1-\eta}\, e^{(\eta-1)\mu s} - (\eta-1)\, g^{-\eta} e^{\left(\eta\mu-\gamma+\frac{\sigma^2}{2}\right)s}. \tag{5}$$

Using equations (2)-(3) and assuming that in the steady state $A_{t+s}/A_t = 1$, the function $f(g,s) = \mathbb{E}_t[\Pi_{t+s}(p,z)/\Pi_t^*(z)]$ gives the expected real static profits evaluated s periods after re-setting the price to a gap g at time t, and scaled by the maximum static real profit at time t. The first term of equation (5) depends on the expected growth in real revenues between t and $t+s$. Notice that, for given firm's price p, the growth rate in real revenues is deterministic because the aggregate price level (and the wage) grows at the constant rate μ. The second term of equation (5) depends on the expected growth in real marginal cost. Notice that the expected growth in marginal cost depends both on expected inflation and expected growth in productivity.

Upon paying the observation cost, the firm finds out the value of productivity and can change its nominal price. The observation costs are proportional to the value of the maximum static monopolist profit at the time the observation takes place, the latter being a random variable:

$$\theta_t(z) = \theta \ \Pi_t^*(z) \ , \tag{6}$$

where $\theta > 0$ is a parameter. The constant of proportionality, θ, is itself a random variable introducing i.i.d. variation in the cost of observing the state. We normalize parameters so that the maximum static monopolist profits equals one when productivity and nominal wages are equal to one. In Appendix E we consider a variation of the model where the frictionless profits, $\Pi_t^*(z)$ are constant so that all the variation of $\theta_t(z)$ stems from variation in θ.

We assume that the draws of observation cost are independent and identically distributed across consecutive reviews. The timing of events is as follows: upon a review the firm learns the value of the (log of) productivity z, learns the value of the *next* observation cost θ, and sets the new price for its good (or equivalently sets g) and the length of time until the next observation τ. The values of θ are i.i.d. across observations (or reviews), and are drawn from a distribution with density $w(\theta)$. To simplify, we have no aggregate variation across firms, i.e the shocks are idiosyncratic.

In our general specification of the model, we assume that price changes can only happen at the time of observation, i.e. we are ruling out *price plans*.[7] This assumption is justified by results in Alvarez, Lippi, and Paciello (2011) where we show that a relatively small adjustment cost makes price plans not optimal in a similar environment. Hence, we interpret the setup of this paper as an approximation of a more general framework where both adjustment and observation costs are present. Later, we will solve the model in the special case of costless price plans in the spirit of Mankiw and Reis (2002).

To better understand the expression of the value function below, we first write the expected discounted nominal profits between times t and $t+\tau$ for a firm that charges the price $p_t = g\, p_t^*$ between t and $t+\tau$:

$$\begin{aligned}
&\int_t^{t+\tau} e^{-(\lambda+\rho+\mu)s}\; W_{t+s}\; \mathbb{E}\left[\, \Pi_{t+s}\left(p_t\,,\; z_{t+s}\right) \mid z_t\right] ds \\
=\;& W_t\; A_t\, z_t^{\eta-1}\; F(1) \int_t^{t+\tau} e^{-(\lambda+\rho+\mu)s}\, \frac{A_{t+s}}{A_t}\, \frac{W_{t+s}}{W_t}\, \mathbb{E}\left[\left(\frac{z_{t+s}}{z_t}\right)^{\eta-1} \frac{F(g_{t+s})}{F(1)} \mid g_t = g\right] ds \\
=\;& W_t\; \Pi_t^*(z_t) \int_t^{t+\tau} e^{-(\lambda+\rho)s}\; \frac{A_{t+s}}{A_t}\; f(g,s)\, ds\;,
\end{aligned}$$

where we use that the nominal interest rate is $\rho+\mu$, that firms die at rate λ, that nominal wages grow at rate μ, as well as the definitions of Π_t, Π_t^* and f. Since the expected discounted profits as well as the observation cost are proportional to $z^{\eta-1}$ we can normalize the value function dividing it by $W_t\; \Pi_t^*$ (which is proportional to $z_t^{\eta-1}$). Hence, the value function of the firm after drawing an observation cost for the next review equal to θ, and scaled by

[7] In the special case with no drift in inflation, $\mu = 0$, and real marginal cost being a martingale, $\gamma = \sigma^2/2$, this assumption is redundant as the optimal price conditional on the information available at the last observation is not changing between observation dates.

$W_t\ \Pi_t^*$, is:

$$v_t(\theta) \quad = \max_{\tau\in\bar{\mathbb{R}}_+, g\geq 0} \int_0^\tau e^{-(\rho+\lambda)s}\ \frac{A_{t+s}}{A_t}\ f(g,s)\,ds\ + \frac{A_{t+\tau}}{A_t}\ e^{-(\rho+\lambda-b)\,\tau}\left[-\theta+\int_{\underline{\theta}}^{\infty} v_{t+\tau}(\tilde{\theta})\,w(\tilde{\theta})\,d\tilde{\theta}\right].$$
$$\text{where} \quad b\equiv(\eta-1)\left(\gamma\ +(\eta-1)\frac{\sigma^2}{2}\right) \tag{7}$$

where $\bar{\mathbb{R}}_+ = \mathbb{R}_+\cup\{+\infty\}$, so never observing the state is feasible. We evaluate the decision of the firm in steady state, i.e. $A_t = A$ for all t, and drop the sub-index for the value function, $v_t(\theta) = v_{t+s}(\theta) = v(\theta)$. Note that the continuation value is discounted at the rate $\rho+\lambda-b$. The first two terms of equation (7) are due to the standard intertemporal discount and the exogenous product mortality. The third term, $b\equiv(\eta-1)(\gamma+(\eta-1)\frac{\sigma^2}{2}) > 0$ is due to the expected value of the frictionless future discounted profits: $\mathbb{E}\left(\Pi_{t+\tau}^*/\Pi_t^*\right)$. Since the frictionless profit is proportional to $z^{\eta-1}$ (see equation (2)), then the expected growth in frictionless profit is higher if η is higher and/or the drift and volatility of the log-normal productivity z increase. For the firm's problem to be well behaved we will require that the discount rate is positive i.e. $\rho+\lambda > b$.

The optimal policy is given by two functions of θ: the length of time until the next observation $\hat{\tau}(\theta)$, and the optimal price gap $\hat{g}(\theta)$. Our notation allows for negative values of θ, i.e. when $\underline{\theta} < 0$. We restrict the parameters in order to have: i) non-negative inflation, i.e. $\mu \geq 0$, ii) a well defined static monopoly problem, i.e. $\eta > 1$, iii) the value of a firm in the frictionless case to be finite, requiring $\lambda+\rho > b$ (as already mentioned), iv) the value of a firm that is never observing and adjusting to be finite, requiring $\lambda+\rho > \max\{\eta\mu-\mu\ ;\ \eta\mu-\gamma+\sigma^2/2\}$.

Given the optimal policy $\hat{\tau}(\theta)$, the distribution of the cost $w(\theta)$, and the (exogenous) product replacement rate λ, we can obtain the implied distribution of times τ elapsed between price observations/adjustments, which we describe by the (right) cumulative distribution function $H(\tau)$. Note that adjustments occur for two reasons: the first one is that the date for observation/adjustment planned by the firm occurs. The second is that the product dies before the planned observation, and a new observation (and adjustment of the price gap) occurs. Therefore, we keep track of two distributions: the right CDF of the times between reviews conditional on the product not being replaced, denoted by the function $\hat{H}(\tau)$, and the unconditional distribution of times between reviews (i.e. inclusive of product replacements), denoted by the (right) distribution function $H(\tau)$. We thus have:

$$\hat{H}(\tau) = \int_{\underline{\theta}}^{\infty} \mathbf{1}_{\{\hat{\tau}(\theta)\geq\tau\}}\,w(\theta)\,d\theta \quad \text{and} \quad H(\tau) = \hat{H}(\tau)\,e^{-\lambda\tau}\ . \tag{8}$$

where the value of $\hat{H}(\tau)$ is obtained by integrating the probabilities according to $w(\theta)$ of the values of the costs that imply optimal times $\hat{\tau}(\theta) \geq \tau$. The distribution $H(\tau)$ is readily obtained by noting that products die at a constant rate λ, triggering new observations.

Next we describe the first order conditions to the firm's problem. It is convenient to define $E_v \equiv \int_{\underline{\theta}}^{\infty} v(\theta)\, w(\theta)\, d\theta$, which measures the expected value of the optimal policy. Given E_v, we show that the solution for $\hat{\tau}$ and $\hat{g}$ is independent of the whole shape of the function $w(\theta)$, so that these decisions will be described by two functions $\hat{\tau}(\theta; E_v)$ and $\hat{g}(\hat{\tau})$. To see this consider the first order conditions with respect to τ and g of the right-hand-side of equation (7):

$$
\begin{aligned}
0 &= \eta\, \hat{g}^{1-\eta}\, e^{(\eta-1)\mu\hat{\tau}} - (\eta-1)\, \hat{g}^{-\eta} e^{\left(\eta\mu-\gamma+\frac{\sigma^2}{2}\right)\hat{\tau}} - (\rho+\lambda-b)\, e^{b\,\hat{\tau}}\ \left[-\theta + E_v\right]\ , \qquad (9)\\
0 &= \int_0^{\hat{\tau}} e^{(-\rho-\lambda+(\eta-1)\mu)s}\, ds - \int_0^{\hat{\tau}} \hat{g}^{-1} e^{\left(-\rho-\lambda+\eta\mu-\gamma+\frac{\sigma^2}{2}\right)s}\, ds \qquad (10)
\end{aligned}
$$

Using equation (10) we can solve for $\hat{g}(\hat{\tau})$ as

$$
\hat{g}(\hat{\tau}) = \frac{1 - e^{\left(-\rho-\lambda+\eta\mu-\gamma+\frac{\sigma^2}{2}\right)\hat{\tau}}}{1 - e^{(-\rho-\lambda+(\eta-1)\mu)\hat{\tau}}}\ \frac{\rho+\lambda-\eta\mu+\mu}{\rho+\lambda-\eta\mu+\gamma-\frac{\sigma^2}{2}} > 0, \qquad (11)
$$

where the strict inequality follows from our assumptions on parameters guaranteeing a well defined firm's problem. Notice that, the larger is expected inflation, μ, relatively to expected growth in productivity, $\gamma - \sigma^2/2$, the larger is the optimal price gap. Intuitively, if the firm expects not to adjust its price for a period of time, it finds optimal to front-load the expected growth in nominal marginal cost upon each adjustment. By replacing $\hat{g}(\hat{\tau})$ in equation (9) we obtain $\hat{\tau}(\theta; E_v)$. Finally, by replacing back $\hat{\tau}(\theta; E_v)$ and $\hat{g}(\hat{\tau}(\theta; E_v))$ into the Bellman equation 7, we obtain one equation in one unknown:

$$
E_v = \frac{\int_{\underline{\theta}}^{\infty} \left(\int_0^{\hat{\tau}(\theta;E_v)} e^{-(\rho+\lambda)s} f\left(\hat{g}\left(\hat{\tau}(\theta;E_v)\right), s\right)\, ds - \theta\, e^{-(\rho+\lambda-b)\hat{\tau}(\theta;E_v)} \right) w(\theta)\, d\theta}{1 - \int_{\underline{\theta}}^{\infty} e^{-(\rho+\lambda-b)\hat{\tau}(\theta;E_v)}\, w(\theta)\, d\theta}\ . \qquad (12)
$$

Equation (12) together with equations (9) and (11) completely characterize the firm's optimal policy.

2.1 General properties of the firm's problem

We next describe some properties of this problem. Some of this properties turn out to be important to later understand the predictions of this model about the distribution of the times until the next review for the cross section of firms and, therefore, about the effects of monetary shocks on output. See Appendix A for the proofs.

Proposition 1. The optimal decision rules and value function satisfy:

1. If $\int_{\underline{\theta}}^{\infty} \theta\, w(\theta)\, d\theta > 0$, then setting $\tau(\theta) = 0$ for all θ is not optimal,
2. $v(\theta)$ is a weakly decreasing function of θ,
3. If $\underline{\theta} \geq 0$, then $\hat{\tau}(\underline{\theta}) > 0$,
4. Let $E_\theta = \int_{\underline{\theta}}^{\infty} \theta\, w(\theta)\, d\theta$ and $\bar{v}$ be the value function for the problem with a degenerate cost distribution concentrated at E_θ. Then $v(E_\theta) \geq \bar{v}$.

In words, Proposition 1 implies that an average observation cost larger than zero is sufficient for having at least some firms not reviewing continuously. The larger the cost of the next observation, the lower the value of expected discounted profits. Interestingly, we find that $\hat{\tau}(0) > 0$, i.e. firms wait some time even when the next observation is free. This is because of the option value of waiting arising from the uncertainty in the draw of future observation costs. In our applications, the option value of waiting will turn out to be relevant to determine the smallest time between reviews for given support of the observation costs. For instance, one important implication is that, in order to endogenously generate a positive mass of arbitrarily small durations of reviews, the distribution of observation costs needs to have positive mass on negative costs, i.e. $\underline{\theta} < 0$ and $\int_{\underline{\theta}}^{0} w(\theta)\, d\theta > 0$.

The next proposition analyzes some comparative statics of the problem as the distribution of observation cost, w, changes.[8] Let Ω be the c.d.f. associated to the density w, and likewise for $\tilde{\Omega}$ and $\tilde{w}$. Let v and $\tilde{v}$ be the value functions corresponding to w and $\tilde{w}$, respectively.

Proposition 2. If $\Omega(\theta) \geq \tilde{\Omega}(\theta)$ for $\theta \geq \underline{\theta}$, then $v \geq \tilde{v}$. Furthermore let $\bar{\Theta}_w$ be the set for which never observing is optimal given density w. If $\Omega(\theta) > \tilde{\Omega}(\theta)$ for all $\theta > \underline{\theta}$, then $v(\theta) > \tilde{v}(\theta)$ for all $\theta \notin \bar{\Theta}_w$.

Proposition 2 shows a natural comparative static result: if the distribution of the cost is stochastically higher, then the value function is everywhere lower, which immediately implies that the expected value function is smaller. This result is useful for the characterization of the optimal policy, since the distribution of the observation cost, w, affects the first order condition for $\hat{\tau}$ only through the expected value of v.

[8] The proofs are omitted since they are straightforward.

2.2 The case of martingale marginal cost or costless price plans

An interesting benchmark case occurs when the $\mu = 0$ and $\gamma = \sigma^2/2$. This benchmark is interesting for two reasons. First, with this parameter configuration there is no inflation, and the expected real marginal cost of production, $1/z_t$, is a martingale, so that the nominal marginal (and average) cost is also a martingale and the discount parameter b defined in equation (7) becomes $b = (\eta - 1)\eta\frac{\sigma^2}{2}$. These assumptions imply that expected real revenues evaluated s periods after setting prices, $f(\hat{g}, s)$, do not depend on s, since inflation is zero, and the real expected cost is constant. Clearly in this case the optimal markup equals the one for the static monopoly so that $\hat{g} = 1$ and, given our normalization, expected real revenues are $f(1, s) = 1$ for all $s \geq 0$. Hence, the firm's problem simplifies to the extent that we can obtain an explicit expression for the firm's decision rule.

Second, we can interpret the solution to the case of $\mu = 0$ and $\gamma = \sigma^2/2$ as the solution to the case where $\mu \neq 0$ and $\gamma = \sigma^2/2$ but *price plans* are allowed at no cost as in Reis (2006). In fact, if price plans are allowed at no cost, the optimal price plan when $\mu \neq 0$ and $\gamma = \sigma^2/2$ will index prices to inflation so that $\hat{g}_s = e^{\mu s}$ is the optimal plan for the price s periods from the observation expressed relatively to the value of p^* at the last observation. At the optimal plan, the firm expects real profits to be constant, so that $f(\hat{g}_s, s) = 1$ for all $s \geq 0$. Therefore, the solution to the firm's problem in the case of costless price plans and inflation is exactly the same of our case with $\mu = 0$ but no price plans.[9]

The Bellman equation for the case of $\mu = 0$ and $\gamma = \sigma^2/2$ becomes:

$$v(\theta) = \max_{\tau \in \bar{\mathbb{R}}_+} \int_0^\tau e^{-(\lambda+\rho)t} dt + e^{-(\rho+\lambda-b)\tau}[E_v - \theta], \tag{13}$$

where as before we assume $\rho + \lambda > b$ so that the value of expected discounted real profits in the frictionless problem is finite. Let $\underline{E}_v$ be the value of expected discounted real profits in the case of never observing, and let $\bar{E}_v$ be the value of expected discounted real profits in the frictionless case, if $\underline{\theta} \geq 0$ we have that E_v is bounded below and above:

$$\underline{E}_v \equiv \frac{1}{\rho + \lambda} \leq E_v \leq \frac{1}{\rho + \lambda - b} \equiv \bar{E}_v \ .$$

The next proposition gives a complete characterization of the optimal policy for this problem.

PROPOSITION 3. Assume that $\rho + \lambda > b$, then:

[9] Note that in the more general case of case of $\mu \neq 0$ and $\gamma \neq \sigma^2/2$, the expected growth in real profits at the optimal plan is given by $f(\hat{g}_s, s) = e^{(\eta-1)(\gamma-\sigma^2/2)s}$. The formulation of the problem is different, but results are similar to the case of a martingale in real marginal cost.

1. The optimal policy is given by:

$$\hat{\tau}(\theta) = \begin{cases} \max\left\{\frac{-\log\left[(\rho+\lambda-b)(E_v-\theta)\right]}{b}, 0\right\} & \text{if } \theta < E_v \\ \infty & \text{if } \theta \geq E_v \end{cases} \tag{14}$$

2. The value of E_v is the smallest solution of the equation:

$$E_v = \frac{1}{\rho+\lambda} + \frac{b}{\rho+\lambda}\left[\rho+\lambda-b\right]^{\frac{\rho+\lambda-b}{b}} \int_{\underline{\theta}}^{E_v} (E_v-\theta)^{\frac{\rho+\lambda}{b}}\, w(\theta)\, d\theta \tag{15}$$

3. The optimal policy at the lower bound for the support for θ satisfies:

$$\hat{\tau}(\underline{\theta}) \begin{cases} > 0 & \text{if } \underline{\theta} \geq 0\,, \\ = 0 & \text{if } \underline{\theta} \leq E_v - \frac{1}{\rho+\lambda-b} < 0\,. \end{cases} \tag{16}$$

4. If $w(\cdot)$ is first order stochastically higher, E_v is smaller, and hence $\hat{\tau}$ is weakly higher pointwise.

5. If $w(\cdot)$ is second order stochastically higher, E_v is larger, and hence $\hat{\tau}$ is weakly lower pointwise.

This proposition shows that $\hat{\tau}(\theta)$ is a simple function, increasing and convex in θ, known in closed form up to one parameter, namely E_v. This proposition also gives an implicit expression for E_v, as well as an extensive analysis of it, with particular emphasis on its dependence on the distribution of cost w. The precise value of E_v and whether there are values for $\theta > E_v$ depends on assumptions about the distribution of cost, w. For instance, if the support of θ is unbounded above, then there must be values for which the optimal policy is to never observe after a draw $\theta > E_v$. Yet, an unbounded support is not necessary for never observing to be optimal.

We conclude by discussing an important property of the function $\hat{\tau}(\theta)$ that will prove useful in modeling the real effects of monetary shocks in Section 4. The objective is to characterize how the variance of transitory changes in observation costs, as summarized by the coefficient of variation $CV(\theta)$, affects the variance of the length of reviews, $\hat{\tau}$. Here we develop an analytical expression for the case of small variance in θ. To this end we use the following approximation for the variance of the function of a random variable: $Var(\hat{\tau}(\theta)) = \left(\hat{\tau}'(\theta)\Big|_{\theta=\bar{\theta}}\right)^2 Var(\theta) + o(Var(\theta))$ where the approximation is taken around the mean value of the observation cost denoted by $\bar{\theta}$, and where $\hat{\tau}'(\theta)$ is the first derivative of the function

$\hat{\tau}(\theta)$ in equation (14). The approximation requires that the coefficient of variation of θ is a strictly positive constant as $\bar{\theta}$ go to zero.[10] Under these assumptions it is readily shown that the equation linking the two (squared) coefficient of variation is (see Appendix B)

$$(CV(\hat{\tau}))^2 = (\rho + \lambda - b)^2 \frac{\bar{\theta}}{2b} (CV(\theta))^2 + o(\bar{\theta}) \tag{17}$$

where b is defined in equation (7). Note that $\bar{\theta}/b$ can be interpreted as the ratio of the mean observation cost to the benefit of reviewing, since the expected value of frictionless profits increases at the rate b. In addition, given the assumption of small variance of θ, the ratio $\bar{\theta}/(2b)$ can be shown to be of the same order of magnitude of $(\mathbb{E}(\hat{\tau})/2)^2$ (see Appendix B), where $\mathbb{E}(\hat{\tau})$ is the average time between reviews. Intuitively, when the ratio of costs to benefits $\bar{\theta}/b$ is smaller, then reviews are relatively more frequent. Everything else being equal, if reviews are more frequent, then the sensitivity of the coefficient of variation of $\hat{\tau}$ with respect to the coefficient of variation of θ is smaller because the continuation value of the firm , E_v, is larger, and the variation in θ is relatively less important to determine the next observation time. Finally, the coefficient $(\rho + \lambda - b)$ is typically of the same order of magnitude of a discount rate, hence relatively small.

An important implication of equation (17) is that, when matching the average frequency of observations and adjustments of modern economies, the sensitivity of the square of coefficient of variation of $\hat{\tau}$ with respect to the square of coefficient of variation of θ is typically very small. For instance, at our baseline parametrization for the U.S. economy, we find that $(\rho + \lambda - b)^2 \frac{\bar{\theta}}{2b} = 0.017$, so that to achieve a given coefficient of variation of $\hat{\tau}$, the model typically needs a coefficient of variation of transitory observation costs which is at least one order of magnitude larger.[11] In Section 4 we will show such small sensitivity matters for the predictions of this model for the real effects of monetary shocks.

3 Steady state aggregation of the firms' decision rules

In this section we characterize the mapping from the distribution of the times between consecutive reviews (implied by optimal firms' decisions), H, to the cross-section distribution of the times until the next review, Q. The last distribution is key to model the propagation of an aggregate shock through the economy as it determines the time it takes before a given

[10] Formally, we consider a sequence of problems for the firm where both the expected value and the variance of the distribution of θ converge to zero, but that they do so keeping the coefficient of variation constant.

[11] In this parametrization, we match a frequency of price changes equal to 1.3 adjustments per year, a discount factor $\rho = 0.02$, and a product substitution rate $\lambda = 0.25$. Finally, we use $b = 0.07$ given a product elasticity $\eta = 4$, and a value of σ that matches the average size of price changes in the U.S. We later discuss these parameters more in detail.

fraction of firms becomes aware and responds to an aggregate shock. Among all the possible distributions, we will pay particular attention to the case of exponential cross-section distribution Q, as this is the distribution of *inattentiveness* assumed in Mankiw and Reis (2002) and in the large applied literature that followed thereafter.[12] We will discuss the extent to which our aggregation results depend on the random observation cost model of Section 2, and the extent to which they hold more generally whenever times between reviews are random, and some other regularity conditions on the distribution H are satisfied.

In the second part of the section we do a different type of exercise: we solve for the distribution of observation cost $w(\theta)$ that produces a given invariant distribution $H(\tau)$, given the other parameters of the model. In fact our model provides a unique mapping between w and H and viceversa. This mapping is useful to characterize what fundamental assumptions are needed to justify e.g. the exponential distribution of *inattentiveness* assumed in Mankiw and Reis (2002), as well as others. In fact, one can interpret the results of this section as the combination of two different mappings. On one side the aggregation results give us a mapping between the distributions Q and H, and on the other side our random observation cost model provides a mapping between w and H.

Throughout this section we assume that there is a measure one of firms facing the same problem but with independent realizations of the processes for productivity and observation costs. The description of the rest of the economy is delayed until Section 4, where we apply the results of this section to study the effects of a monetary shock.

3.1 The cross-sectional distribution of times until the next review

In this section we take the optimal policy of a firm in the form of the distribution H, and derive the implication for the cross sectional distribution Q. The advantage of using H as a primitive is that it will allow us on one side to compare to existing literature (e.g. Reis (2006)), and on the other side to interpret the model with random observation costs of Section 2 as a microfoundation for a given H.

Using the same notation of Section 2, let $H : \mathbb{R}_+ \to [0, 1]$ denote the (time invariant) right CDF of the time elapsed between consecutive reviews, a weakly decreasing function. We assume: i) that the times elapsed between consecutive reviews for each firm are identically and independently distributed, ii) that H has at most countably many downward jumps, and iii) that H is continuous at $\tau = 0$. Furthermore we assume iv) that there is a unit mass of firms, each of which draws times for the next review from H independently, and that a law of large number applies to them. In words, $H(\tau)$ is the fraction of firms that upon conducting

[12] See Mankiw and Reis (2010) for a review of the literature.

a review "decide" to conduct their next review after τ periods or more.[13] Notice that the model of Section 2, and in particular firms' optimal decision rule given by equations (25)-(12), satisfies assumptions i) to iv), hence the model is a microfoundation.

We now consider the distribution across firms at a given calendar date t of the times until the next review. We denote by $Q_t(r)$ the fraction of firms that at date t will conduct their next review r periods or more from now (i.e. no sooner than date $t+r$). Since we are seeking an invariant distribution we drop the calendar time t and refer to Q as the distribution of the times until the next review. The function $Q : \mathbb{R}_+ \to [0, 1]$ is also a right CDF, and hence weakly decreasing. We denote by $q(r) \equiv -Q'(r)$ the density of the time until the next review.

Proposition 4. *If H satisfies assumptions i) to iv), then the density of the invariant distribution of the times r until the next review is*

$$q(r) = \frac{H(r)}{\int_0^\infty H(s)\,ds} \text{ for all } r \geq 0 \ . \tag{18}$$

A proof is given in Cox and Miller (2001), Heyman and Sobel (2003) or in Appendix A.[14] An immediate implication of Proposition 4 is given in the next corollary:

Corollary 1. *The cross section distribution of the time until the next review, Q, is exponential if and only if the firm's distribution of the times between reviews, H is exponential.*

The proof follows since an exponential density is proportional to its right CDF, a property shared by $q(r) \propto H(r) = e^{-\lambda r}$ in equation (18) for $\lambda > 0$. Notice that these results contradict Reis (2006), where under the same assumptions i) to iv) described above, it is incorrectly claimed that the cross section distribution of the time until the next review, Q, is exponential *for any* distribution H. We provide more details on how our results relate to Reis (2006) in Section 5.

Next we use two analytical examples to illustrate the workings of Proposition 4.

[13] The resulting series of reviews form a stationary renewal process, see Cox and Miller (2001), page 340.

[14] This result can be found in Cox and Miller's (2001) equation (29) on page 347, or equation (52) on page 356. Using Cox and Miller's (2001) notation, q is the limiting or stationary density of the forward recurrence times (see page 347). Alternatively, see part (b) of Theorem 5.10 in Heyman and Sobel (2003) on page 133, where it is shown that equation (18) gives the "equilibrium excess distribution" of the forward recurrence times, i.e. what we denote as the invariant distribution of times until the next review.

Example 1. Assume the right CDF of the time between consecutive reviews, $H(\tau)$, is

$$H(\tau) = \begin{cases} e^{-\lambda \tau} & \text{if } \tau \in [0, \bar{\tau}] \\ 0 & \text{if } \tau \in (\bar{\tau}, \infty) \end{cases} \tag{19}$$

then the cross section density of the times r to the next review is

$$q(r) = \begin{cases} \frac{\lambda e^{-\lambda r}}{1 - e^{-\lambda \bar{\tau}}} & \text{if } r \in [0, \bar{\tau}] \\ 0 & \text{if } r \in (\bar{\tau}, \infty) \end{cases} \tag{20}$$

In this example the distribution H has two parameters, $\bar{\tau}$ and λ. This corresponds to the times between reviews τ being drawn from a distribution with a mass point at $\bar{\tau}$, i.e. the probability that $\tau = \bar{\tau}$ is $e^{-\lambda \bar{\tau}}$, and for $\tau \in [0, \bar{\tau})$ the times are drawn from an exponential density decaying at rate λ. This example can be used to discuss two important cases. First, the case where $\lambda = 0$ and $\bar{\tau} > 0$ corresponds to the case where each firm reviews exactly every $\bar{\tau}$ periods, i.e. the distribution $H(\tau)$ is degenerate at $\bar{\tau}$. This case corresponds to the optimal decision rules derived by in Caballero (1989) and Section 4 in Reis (2006), as well as in our model of Section 2 when there are no substitution, $\lambda = 0$, and the distribution of cost w is degenerate with all mass at one level of θ. The cross sectional distribution $q(r)$ for this case is the uniform distribution with support in $[0, \bar{\tau}]$. The second case is obtained when $\bar{\tau} \to \infty$ and $\lambda > 0$. In this case the distribution $H(\tau)$ is exponential, and so is the cross sectional distribution $q(r)$. This case corresponds to the assumptions made in Sections 5 and 6 of Reis (2006) and in Mankiw and Reis (2002), and corresponds to a specific assumption about the distribution of observation cost w in our model of Section 2, which we characterize in the next Section.

For completeness we give another simple example. In this example the distribution $H(\tau)$ represents times drawn from a uniform density over the support $[0, \bar{\tau}]$. The implied cross section distribution $q(r)$ has a linearly decaying density.

Example 2. Assume the right CDF of the time between reviews, $H(\tau)$, is

$$H(\tau) = \begin{cases} 1 - \frac{\tau}{\bar{\tau}} & \text{if } \tau \in [0, \bar{\tau}] \\ 0 & \text{if } \tau \in (\bar{\tau}, \infty) \end{cases} \tag{21}$$

then the cross section density of the times r to the next review is:

$$q(r) = \begin{cases} \frac{1-r/\bar{\tau}}{\bar{\tau}/2} & \text{if } r \in [0, \bar{\tau}] \\ 0 & \text{if } r \in (\bar{\tau}, \infty) \end{cases} \tag{22}$$

3.2 Rationalizing a distribution of times between reviews

In this section we solve a "reverse engineering problem": we derive the distribution of the observation costs $w(\theta)$ that is consistent with a given target for the distribution of time elapsed between observations (conditional on the product not being replaced) $\hat{H}$, given the other parameters of the model $(\eta, \rho, \sigma^2, \mu, \gamma)$.[15] This exercise is useful because we can evaluate the qualitative and quantitative properties of the distribution of observation costs that would support a given $\hat{H}$ and, given the results of the previous section, a given Q.

Assuming that $\hat{\tau}(\theta; E_v)$ is increasing in θ, and denoting the cumulative distribution function of observation cost by $\Omega(\theta) = \int_{\underline{\theta}}^{\theta} w(\theta)\, d\theta$, we have $\Omega\left(\hat{\tau}^{-1}(\theta; E_v)\right) = 1 - \hat{H}(\hat{\tau})$. Finally, using $\hat{\tau}(\theta; E_v)$ from equations (9)-(10) together with E_v in equation (12), we obtain one equation in one unknown, E_v. Denoting by E_v^* the solution of equation (12), and abusing notation, the optimal policy is given by $\hat{\tau}(\theta) = \hat{\tau}(\theta; E_v^*)$. By differentiating $1 - \hat{H}(\hat{\tau})$ with respect to θ, the implied density of observation cost evaluated at θ is given by

$$w(\theta) = -\hat{H}'\left(\hat{\tau}(\theta)\right) \frac{\partial \hat{\tau}(\theta)}{\partial \theta} \quad \text{for all } \theta > \underline{\theta}\ . \tag{23}$$

We can use equation (23) to "reverse engineer" w from any given $\hat{H}$, and study its qualitative and quantitative properties. In the next section we apply this result to the case of exponential $\hat{H}$, because this is of particular interest given the existing literature. In Section 4.1 we will consider alternative assumptions for $\hat{H}$.

3.2.1 Rationalizing an exponential distribution of times between reviews

Of particular interest, given results by Mankiw and Reis (2002), is to use equation (23) to characterize the distribution of $w(\theta)$ that gives $\hat{H}(\tau) = \exp\left(-\xi\tau\right)$ for some $\xi > 0$ and all $\tau > 0$. Equation (25) readily implies that if $\hat{H}$ is exponential then H is exponential,

[15] Recall the assumption that products are replaced at rate λ per unit of time, triggering a new review for the replaced product, so that the combined time between consecutive reviews has right CDF equal to $H(\tau) = \hat{H}(\tau)\exp(-\lambda\tau)$. For the purpose of this section we are going to use $\hat{H}(\tau)$ as a "primitive". Transformation from this function to $H(\tau)$ is straightforward.

$H(\tau) = \exp(-(\xi+\lambda)\tau)$, so that Corollary 1 immediately implies that the cross section distribution Q is also exponential.

First, consider the case with nominal marginal cost being a martingale (or equivalently to the case where we allow for costless price plans). For this case we can produce a closed form solution, up to a location parameter of the density $w(\theta)$, so that the distribution $\hat{H}(\tau)$ is an exponential with parameter ξ.

Proposition 5. If $\mu = 0$ and $\gamma = \sigma^2/2$, or if price plans are allowed at no cost and $\gamma = \sigma^2/2$, the distribution of random observation cost leading to an exponential distribution of times between reviews, $\hat{H}(\tau) = \exp(-\xi\tau)$ with $\xi > 0$, has the form:

$$w(\theta) \quad \propto \quad [E_v - \theta]^{\frac{\xi}{b}-1} \quad \text{for } \theta \in (\underline{\theta}, E_v) \ , \tag{24}$$

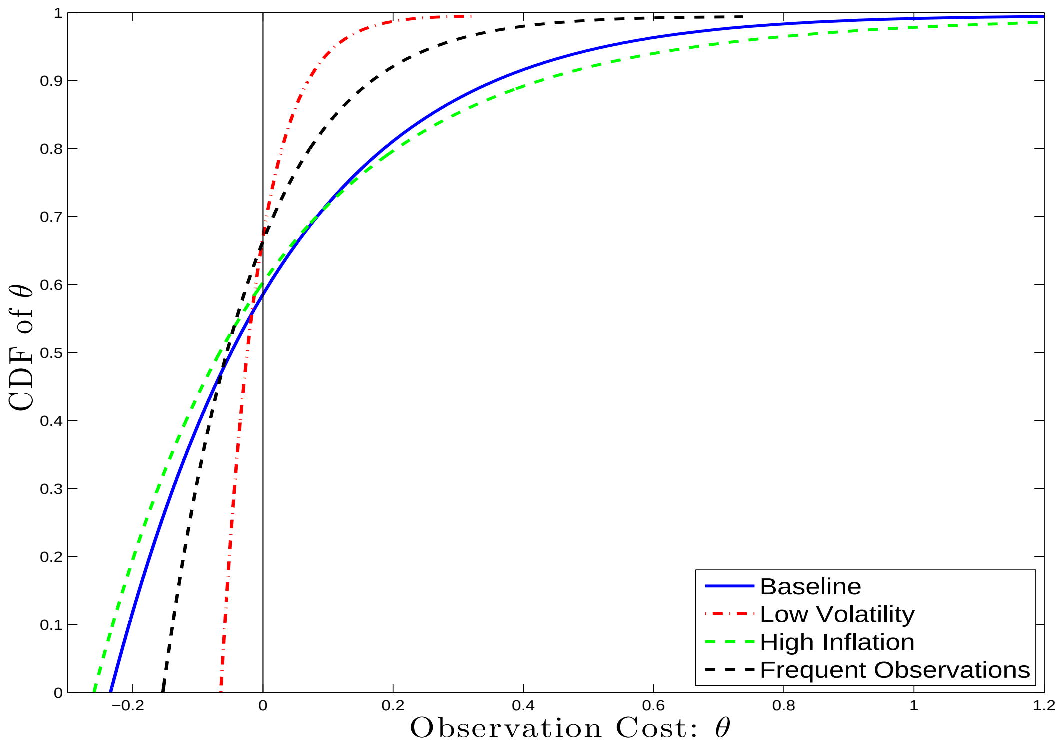
with $\int_{\underline{\theta}}^{0} w(\theta) d\theta > 0$.

The proof follows immediately from using equation (23) and results in Proposition 3. In particular, the fact that $\int_{\underline{\theta}}^{0} w(\theta) d\theta > 0$ follows from the fact that $\hat{\tau}(\theta) > 0$ at $\theta = 0$, that $\hat{\tau}'(0) > 0$, and that the exponential distribution assigns a positive mass to $\hat{\tau}$ arbitrarily close to zero. Hence, in order to support exponential distributed times as in Mankiw and Reis (2002), the random observation cost model needs strictly positive mass on negative observation costs.

Next, we study the more general case where the marginal cost is not a martingale and price plans are costly. We derive numerically the distribution of random observation cost supporting an exponential distribution of reviews H, arrival rate of reviews equal to ξ, by using $\hat{\tau}(\theta; E_v)$ from equations (9)-(10) together with $w(\theta)$ in equation (23). We choose the remaining parameters of the model to match U.S. CPI statistics about timing and size of price changes estimated by Nakamura and Steinsson (2008). The parameter ξ associated to the exponential distribution $\hat{H}$ is completely determined by matching the average frequency of price adjustments equal to 1.3 adjustments per year. The frequency of product substitutions is assumed to be equal to $\lambda = 0.25$. The annual discount rate is $\rho = 0.02$. We assume a steady state markup of one third implying $\eta = 4$. We choose $\sigma = 0.1$, so that the standard deviation of log-price changes is equal to 0.087. Finally, we assume $\mu = 0$ and $\gamma = \sigma^2/2$ so that we have an analytical characterization of our benchmark case given in equation (24), and we can study how w changes when we deviate from this benchmark.

We evaluate the quantitative properties of the distribution $w(\theta)$ that follows our baseline parametrization, as well as studying how such density varies as a function of the parameters of the model. Figure 1 shows the cumulative distribution function associated to the density of observation costs w in the benchmark case of no drift, in the case with lower volatility

Figure 1: CDF of θ producing an exponential distribution of adjustment times $\hat{H}(\tau)$



Note: The figures plots the CDF of θ: $\int_{\underline{\theta}}^{\theta} w(\theta)\, d\theta$, where observation costs θ are measured in proportion to yearly frictionless profits. In the baseline parametrization the density of observation costs implies an invariant distribution $\hat{H}(\tau)$ that is exponential with $\xi = 1.05$ observations/price changes per year. Since there are also $\lambda = 0.25$ substitutions per year, the baseline case has 1.3 price changes per year. Other parameters of the model are $\eta = 4$, $\sigma = 0.10$, $\gamma = \sigma^2/2$, $\lambda = 0.25$, $\rho = 0.02$. The "Low Volatility" case has $\sigma = 0.05$. The "High Inflation" case has $\mu = 0.05$. The "Frequent Observations" case has $\xi = 2.00$.

$\sigma = 5\%$, in the case with higher drift $\mu = 5\%$, and in the case with higher average frequency of observations, $\xi = 2.00$. We draw the following conclusions. First, in all cases we study, the distribution w is such that more than 50% of mass is on observation costs smaller than zero, with the smallest value being between -5% and -25% of the maximum static monopolist profit per year. The economic interpretation of this result is that observation "benefits" as well as observation costs are needed to produce an exponential distribution of times. Second, there is a large variability of observation costs, with a coefficient of variation equal to $CV(\theta) = 11$, and positive mass on observation cost that are very large, for instance larger than 100% of the maximum static monopolist profit per year. At our baseline parametrization, the average cost of a review is about 2.2% of profits, while the standard deviation is about 24% of profits. The need of large variability in observation cost is explained by our finding in equation (17). Third, the cases with higher frequency of observations/adjustments, or lower volatility, support a distribution of observation cost which is more concentrated around the mean value. Fourth, the distribution is not very sensitive to the level of the drift and/or to allowing for price plans.[16]

We conclude that an exponential distribution of times between reviews H and, therefore, an exponential cross section distribution Q, can be supported by random idiosyncratic observation cost, but only allowing for enough variability in such costs, inducing a positive mass on negative values and allowing for relatively large observation cost as a share of profits. The need of negative observation costs is explained by the option value of waiting associated to the decision of when having the next review even if the next review is costless, which we analytically characterized in equation (14), together with the need of a strictly positive mass of relatively small durations of reviews that characterize the exponential distribution. At our baseline parametrization, such mass on negative values of θ is substantial. Moreover, the model also needs substantial variability in observation cost. The need of a large variability in observation cost is a more general property of the model with random observation cost, and it is not special to matching the exponential $\hat{H}$. In fact, the exponential distribution has a coefficient of variation equal to one, and according to equation (17) the model needs large variability of θ to match it, to the point that the standard deviation of observation costs is more than ten times its mean. However, we could have targeted a different distribution $\hat{H}$ (as we do in Section 6), but as long as such distribution implies a coefficient of variation similar to the exponential, the need of large variability of θ would remain. This is important as in Section 4 we argue that such variability of durations of reviews is an important factor in generating large effects of monetary shocks on output.

3.3 The cross-sectional distribution of price changes

The previous sections studied the predictions of the model for the cross-section distribution of times until the review, Q. This section presents the model implications for the cross-sectional distribution of price changes, and in particular for the variance and kurtosis of price changes.

Proposition 6. The cross-section invariant distribution of log-price changes, $\Delta \log(p)$, is given by a mixture of normals indexed by τ, where the mixture has density $h(\tau) = -H'(\tau)$, and each of the normals has mean and variance $((\mu - \gamma)\,\tau\,,\,\sigma^2\,\tau)$. The variance of the

[16] Appendix E shows that very similar results are produced by a model with correlated shocks to demand and marginal costs, where the π_t^* is constant so that variability in the level of observation costs only come from randomness in θ.

log-price changes is

$$Var(\Delta \log(p)) = \frac{\sigma^2}{N_a} \tag{25}$$

where $N_a \equiv 1/\int_0^\infty H(s)\, ds$ is the average frequency of price changes, and the kurtosis of the log-price changes is equal to:

$$Kurt(\Delta \log(p)) = 3\ [CV(\tau)^2 + 1] \tag{26}$$

where $CV(\tau)$ is the coefficient of variation of the times to review implied by $H(\tau)$.

Proposition 6 illustrates a simple way to calibrate the model to the data, as the parameter σ is pinned down by the observations on the frequency of price changes and the standard deviation of price changes. Moreover equation (26) states that the kurtosis of the distribution of price changes is solely a function of the coefficient of variation of the times to the next observation. For instance if τ has a degenerate distribution, as in the models of Caballero (1989), Bonomo and Carvalho (2004) and section 4 of Reis (2006), prices changes are normally distributed, and hence have kurtosis equal to 3. Higher dispersion in τ, as in the model we develop in Section 2, implies a kurtosis larger than 3. Recall that, at least for a symmetric unimodal distribution, its kurtosis measures the presence of small as well as large observations. We think that the kurtosis of the price changes is an interesting statistic because the first generation fixed cost models such as Golosov and Lucas's (2007) imply very small values – close to one– while the inattentive model implies much larger values. Estimates of the kurtosis of price changes using micro data are typically larger than the ones corresponding to a normal distribution.[17]

4 Impulse response to a monetary shock

In this section we describe the relation between a given distribution of times to review, H, and the impulse response of the aggregate price level and output to a permanent unexpected increase in the money supply. We do so under the assumption that price changes occur only at the times of reviews. We study how the real effects of a monetary shock depend on the

[17] The estimates of the kurtosis are sensitive to measurement error and vary across data sets, but are typically larger than the ones corresponding to normal distributions. Klenow and Malin (2010) and Bhattarai and Schoenle (2010) find kurtosis of 10 or higher using Consumer and Producer price indexes respectively for the US. Midrigan (2011) finds a kurtosis around 4 in scanner data after removing product specific mean and variances. In Alvarez, Le Bihan, and Lippi (2012) we use data from the French CPI and find a kurtosis on raw data higher than 10 (around 10 in standardized data). Alberto Cavallo has generously computed kurtosis for French online stores and found figures similar to ours. Eichenbaum et al. (2012) argue that the small price changes that are seen in the data could be due to measurement error, and hence it will affect, and likely reduce, the measure of kurtosis.

mean and the variability of the durations of price spells between reviews.

The background is given by the DSGE model described in Alvarez, Lippi, and Paciello (2012), which is a variation of Golosov and Lucas (2007). There are two types of agents making decisions in this economy, a representative household and a unitary mass of monopolistically competitive firms. The problem of the monopolistically competitive firms is described in Section 2. The solution to the firms' problem is given by the optimal time between consecutive reviews, $\hat{\tau}$, and optimal price gap upon a review, $\hat{g}$, satisfying equations (9)-(12).

The household has additive time separable preferences, discounted at rate ρ. Each period utility is linear in leisure, logarithmic in real balances, and has CRRA utility with parameter ϵ in aggregate consumption, itself a Dixit-Stiglitz aggregate of goods produced by each firm with elasticity of substitution η. More in detail, the representative household has preferences given by

$$\int_0^\infty e^{-\rho t} \left(u\left(c_t\right) - \alpha L_t + \log \frac{\hat{m}_t}{P_t} \right) dt \quad \text{where} \quad c_t = \left(\int_0^1 C_t(i)^{\frac{\eta-1}{\eta}} \right)^{\frac{\eta}{\eta-1}}$$

where $u(c) = (c^{1-\epsilon} - 1)/(1-\epsilon)$, $C_t(i)$ is the consumption of product i, L_t are labor services, $\hat{m}_t$ is the nominal quantity of money demand, P_t is the nominal ideal price index of one unit of aggregate consumption, and $\rho > 0, \epsilon \geq 1, \alpha > 0, \eta > 1$ are parameters. The elasticity of substitution between any two products η is the same, regardless of the firms that produced them. The stock of money supply, m_t, grows at a rate μ.

The equilibrium level of aggregate consumption is given by

$$c_t = \left(\int \left(\frac{\eta}{\eta - 1} \frac{g}{z} \right)^{1-\eta} \phi_t(dg, dz) \right)^{\frac{1}{\epsilon\,(\eta-1)}}, \tag{27}$$

where $\phi_t(\cdot)$ is the distribution of firms over price gaps and productivity, whereas the equilibrium price level clears the market for real balances,

$$\frac{m_t}{P_t} = (c_t)^\epsilon \ , \tag{28}$$

so that $\hat{m}_t = m_t$. An equilibrium consists of a fixed point in sequences $\{c_t\}_{t\geq 0}$. In Alvarez, Lippi, and Paciello (2012) we show that there exists an equilibrium for this economy where aggregate consumption is constant, the nominal wage W_t, and the price level P_t, grow at the constant rate μ, and the nominal interest rate is constant: $W_t = \alpha\ R_t\ m_t$ and $R_t = \rho + \mu$.

We study an economy that starts at the invariant distribution of firm's price gaps, g, $\phi(dg, dz)$. We assume that at time $t = 0$ there is an unanticipated permanent increase in the level of the money supply by δ log points that causes the distribution of firm's price gaps, g, to shift. We use the result that, for small shocks δ and small observation cost θ, the general equilibrium feedback effect on firms' decision rule is negligible, so that ignoring the effect of the shock on the firm's decision rules (e.g. on the value of $\hat{\tau}$) gives a good approximation of the firm's behavior during the convergence to the steady state.[18]

We study the effect of an aggregate monetary shock of size δ on the deviations of the aggregate price level $P(t)$ at $t \geq 0$ from the initial level $P(0)$, denoted by $\mathcal{P}(\delta, t) = P(t)/P(0)$. In addition, let $\mathcal{P}(\delta, t; r \geq t)$ and $\mathcal{P}(\delta, t; r < t)$ be the deviations from $P(0)$ of the average of the price levels at $t \geq 0$, among the subset of firms with age of last review, r, larger than t and smaller than t, respectively (all averages use equal weights). It follows from the definition of $\mathcal{P}(\delta, t)$ that $\mathcal{P}(\delta, t) = Q(t)\ \mathcal{P}(\delta, t; r \geq t) + (1 - Q(t))\ \mathcal{P}(\delta, t; r < t)$, where $Q(t)$ is the fraction of firms that have made their last review more than t periods before, i.e. the distribution of *inattentiveness* discussed in Section 3.[19]

Let us now evaluate $\mathcal{P}(\delta, t; r \geq t)$ and $\mathcal{P}(\delta, t; r \geq t)$. To simplify the discussion consider first the case where $\mu = 0$ and $\gamma = \sigma^2/2$. However, note that the results of this section about the impulse response of output to the monetary shock do not depend on this assumption. After the monetary shock of $\delta > 0$ log points, nominal wages jump up by this amount and stay constant, while the other nominal quantities will converge to values δ log points higher. A firm that reviews its price for the first time after the aggregate shock will adjust its price to the profit maximizing price (because of the nominal marginal cost is a martingale), which has increased by a factor $\exp(\delta - \gamma\tau - \sqrt{\tau}\sigma s)$, where τ is the time elapsed since the last review and where s is a standard normal innovation in its cost. Averaging across all such firms we obtain that $\mathcal{P}(\delta, t; r < t) = \exp(\delta)$. Similarly, averaging across all firms that have not yet made a review since the monetary shock yields $\mathcal{P}(\delta, t; r \geq t) = 1$. We can thus express the impulse response of the price level t periods after the shock as $\log \mathcal{P}(\delta, t) \ \approx \hat{\mathcal{P}}(\delta, t) = \delta\,(1 - Q(t))$.[20]

Proposition 4 gives an expression for the right cumulative distribution of the times until the next review, $Q(\cdot)$, in terms of the right cumulative distribution of time elapsed between reviews $H(\tau)$. Thus we can express the impulse response of the price level in terms of the

[18] In Alvarez, Lippi, and Paciello (2012) we solve numerically various general equilibrium versions of this model (including menu costs or observation costs or both) and verify the quality of the approximation when the feedback effect is ignored. In Alvarez and Lippi (2012) we study analytically the order of the approximation of neglecting the feedback effect for a model which corresponds to the version in this paper of the firm's problem with perfectly correlated idiosyncratic demand and cost shocks outlined in Appendix E.

[19] Cox and Miller's (2001) equation (52) on page 356 gives the density of the forward recurrence time, but it is remarked that it equals the one of the backward recurrence time, or the age of the reviews. These two distributions are identical.

[20] The approximation is accurate for small δ since $\log\left(1 + (e^{\delta} - 1)\,(1 - Q(t))\right) = \delta\,(1 - Q(t)) + o(\delta)$.

right cumulative distribution of the time elapsed between reviews $H(\tau)$:

$$\hat{\mathcal{P}}(\delta, t) = \delta \left[1 - \frac{\int_t^{\infty} H(s)\, ds}{\int_0^{\infty} H(\tilde{s})\, d\tilde{s}} \right] . \tag{29}$$

Similarly to the price level, we can compute the output response to the monetary shock using the distribution $H(\tau)$. Output in this model equals aggregate consumption, $Y_t = c_t$, and given the path of the aggregate price level and money supply can be obtained from equation (28). Let $\mathcal{Y}(\delta, t) = Y_t/Y_0$ be the deviation of output evaluate t periods after the monetary shock of size δ from the initial level Y_0, which corresponds to the steady state level of output. Equations (28) and (29) imply $\log \mathcal{Y}(\delta, t) \approx \hat{\mathcal{Y}}(\delta, t) = (1/\epsilon)(\delta - \hat{\mathcal{P}}(\delta, t))$, so that:

$$\hat{\mathcal{Y}}(\delta, t) = \frac{\delta}{\epsilon} \frac{\int_t^{\infty} H(s)\, ds}{\int_0^{\infty} H(\tilde{s})\, d\tilde{s}} \quad . \tag{30}$$

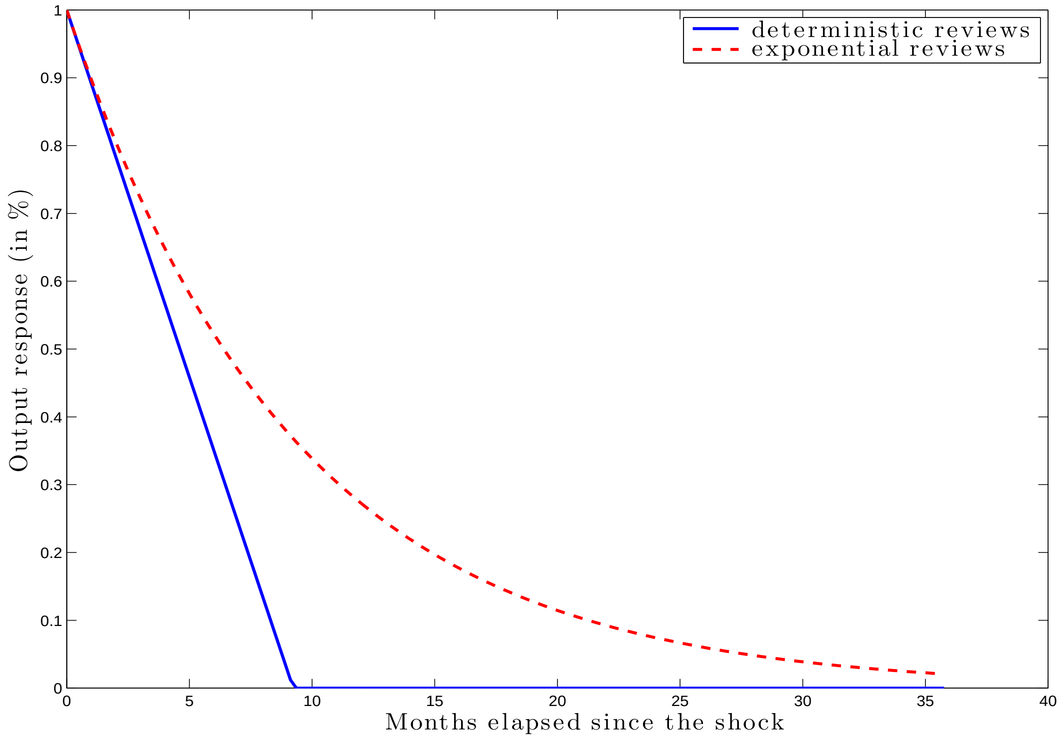
Given the properties of $H(\cdot)$ we note that the impulse response of output must be positive and decreasing. Figure 2 uses equation (30) to plot two impulse responses corresponding to deterministic reviews every $1/N_a$ periods and an exponential distribution H *with the same* average time between reviews. The shock size is set to 1 percent ($\delta = 0.01$) and, for simplicity, the figure furthermore uses $N_a = 1.3$, $\lambda = 0$ and $\epsilon = 1$. Different parameters choices would simply scale the shape of the impulse response displayed in the figure, as shown by equation (30). The profile of the impulse response corresponding to the deterministic reviews has a linear shape, typical of the staggered mechanisms a la Taylor. It appears that the output effect in this case is uniformly smaller than the one obtained when the distribution of the times between reviews is exponential.

Next we construct a summary measure of the effect of money on output is the area below the impulse response function, whose approximation we denote as $\mathcal{M}(\delta)$:

$$\mathcal{M}(\delta) \equiv \int_0^{\infty} \hat{\mathcal{Y}}(\delta, t) dt \approx \int_0^{\infty} \log \frac{Y_t}{Y_0}\, dt \quad . \tag{31}$$

An advantage of this measure is that it takes into account the impact effect as in Caballero and Engel (2007) as well as the persistence of the effect. Another advantage is that it is easy to characterize this magnitude as a function of the structural parameters, as we have done in Alvarez and Lippi (2012) in the analysis of menu cost models, including Taylor-type of staggered adjustments. The next proposition characterizes the area under the impulse response function for output in terms of two moments of the distribution H (see Appendix A for a proof).

Figure 2: Output response to a once and for all monetary shock



Note: Output response to a 1 percent shock to money ($\delta = 0.01$) under 2 distributions of the times to review: a degenerate one with reviews every $1/N_a$ periods and an exponential one with the same average time between reviews (the $H(t)$ expressions appear in Table 1).

Proposition 7. Fix a distribution of times elapsed between reviews/adjustments, $H(\tau)$, and let $CV(\tau)$ be its coefficient of variation, $\mathbb{E}(\tau)$ its expected value, and $N_a = 1/\mathbb{E}(\tau)$ the expected number of adjustments per unit of time. Then the area under the impulse response of output for a monetary shock of size δ is given by:

$$\mathcal{M}(\delta) = \frac{\delta}{\epsilon} \frac{\int_0^\infty s\, H(s)\, ds}{\int_0^\infty H(\tilde{s})\, d\tilde{s}} = \frac{\delta}{\epsilon} \frac{1 + CV(\tau)^2}{2\, N_a} \quad . \tag{32}$$

Proposition 7 states that, keeping fixed the average time between observations/price changes, the higher the variability of time between observations/price changes, the larger the cumulative effect of monetary shocks, δ, on output. In fact, the first expression in equation (32) implies that keeping the same expected value of times between reviews, $\mathbb{E}(\tau) = \int_0^\infty H(\tilde{s})\, d\tilde{s}$, the cumulated effect of a given shock on output is greater for distributions $H(\tau)$ that have more mass on large values, i.e. a larger coefficient of variation. Our Proposition 7 is related to recent results by Carvalho and Schwartzman (2012). Taking as given $H(\cdot)$ they develop an aggregation result for given policy rules and a more general process for the money supply.[21] We focus on a once and for all shock because in the context of our model the firm's decision rules are optimal.

We illustrate Proposition 7 by showing how different distributions of times between reviews, $H(\tau)$, map into the cumulative output response to monetary shocks. These examples show how important the variability of τ is for the cumulative response of output to monetary shocks. The results from this first application are general in the sense that they do not depend on a specific model of information acquisition decision generating a given $H(\tau)$. The second group of applications, discussed in Section 4.1, focuses instead on the model of endogenous information acquisition decision of this paper, and studies how different distribution of random observation cost imply different variability of the time between reviews, i.e. different coefficient of variation. The latter is interesting because while τ and its distribution are endogenous to the model, the distribution of the observation cost is a primitive of our model. Hence, we are interested in studying how much (cumulative) output response to a monetary shock our model can generate as a function of different assumptions about the distribution of the observation cost.

Let us first consider three specifications for $H(\tau)$. Let assume that all three specifications are such that the average frequency of observation, N_a, is the same.

Table 1: Cumulated output response for various $H(\tau)$ distributions

CDF of time between reviews τ	Cumulated Output Response, $\mathcal{M}(\delta)$
Degenerate: $H(s) = 1$ for $s \in [0, \frac{1}{N_a}]$, $H(s) = 0$ for $s > \frac{1}{N_a}$	$\frac{\delta}{\epsilon} \frac{1}{2N_a}$
Exponential: $H(s) = \exp(-s\ N_a)$ for all $s \geq 0$	$\frac{\delta}{\epsilon} \frac{1}{N_a}$
Binary: $H(s) = \frac{1}{\bar{\tau} N_a}$ for $s \in (0, \bar{\tau}]$, $H(s) = 0$ for $s > \bar{\tau}$	$\frac{\delta}{\epsilon} \frac{\bar{\tau}}{2N_a}$

In all cases the average frequency of observations/adjustments is N_a.

In the first case reviews have a degenerate distribution with $\tau = 1/N_a$, so that $H(s) = 1$

[21] Their results imply an equation that is the same of our equation (32) when applied to our case of a once and for all shock to the money supply.

for $s \leq 1/N_a$ and zero otherwise. This case is interesting because it coincides with the models studied by the Caballero (1989), Bonomo and Carvalho (2004) and Reis (2006). In the second case, we assume exponential price reviews/changes, so that $H(s) = \exp(-N_a\ s)$. This case is interesting because it coincides with the model studied by Mankiw and Reis (2002), and resembles *Calvo* model of price adjustments. In the third case, we assume that the time between reviews can take two values: 0 or $\bar{\tau} \geq 1/N_a$, the latter with probability $1/(\bar{\tau} N_a)$.[22] Table 1 reports the cumulated output responses in these three different cases computed by using equation (32). According to equation (32), for any given N_a, the distribution of times between reviews that yields the lowest cumulated output response is the one that has zero variance in such times, i.e. $CV(\tau) = 0$. This is the case for the constant observation cost model studied by Caballero (1989), Bonomo and Carvalho (2004) and Reis (2006), which implies a degenerate distribution of τ. The exponentially distributed case studied by Mankiw and Reis (2002) predicts a cumulated output response that is twice as large as the one predicted by the model with constant observation cost. This is true for all parameter values. Finally, in the third example the cumulated effect can be arbitrarily large since the distribution of $H(\tau)$ has a right tail that can be made as large as desired by setting a large value of $\bar{\tau}$. While this distribution has a discontinuity (i.e. $\tau = 0$ happens with strictly positive probability), a continuous approximation to $H(\tau)$ would have similar implications.

We finish this section with a comment on the relationship between the effect on output of a monetary shocks, summarized by $\mathcal{M}(\delta)$, and properties of the frequency and kurtosis of price changes. Replacing equation (26) into equation (32) gives:

$$\mathcal{M}(\delta) = \frac{\delta}{\epsilon} \frac{Kurt\,(\Delta \log(p))}{6\, N_a} \,. \tag{33}$$

Thus if the model were to be consistent with the kurtosis of individual firm's price changes, the area under the impulse response of the output to a monetary shock should satisfy equation (33). For instance, choosing $Kurt\,(\Delta \log(p)) \approx 5$, $N_a \approx 1.3$ to match U.S. statistics (as discussed above) and $\epsilon = 1$ (log utility), one obtains $\mathcal{M}(\delta) = 0.64\,\delta$. Note that this is a relatively large effect of monetary policy shocks, i.e. a once and for all 1% increase in money supply has a cumulated output response that is approximately equivalent to a 0.64 % increase in output for a year.

[22] We consider this case because of its simplicity. Formally this can be seen as the limiting case of a model where there are 2 strictly positive times between reviews, $\underline{\tau}, \bar{\tau}$, satisfying $0 < \underline{\tau} < 1/N_a < \bar{\tau}$. The probability with which $\bar{\tau}$ is C/N_a where $C = \frac{1 - N_a \underline{\tau}}{\bar{\tau} - \underline{\tau}}$ which implies that the mean expected time is $1/N_a$, as in the other models. Letting $\underline{\tau} \downarrow 0$ gives the specification used in the table.

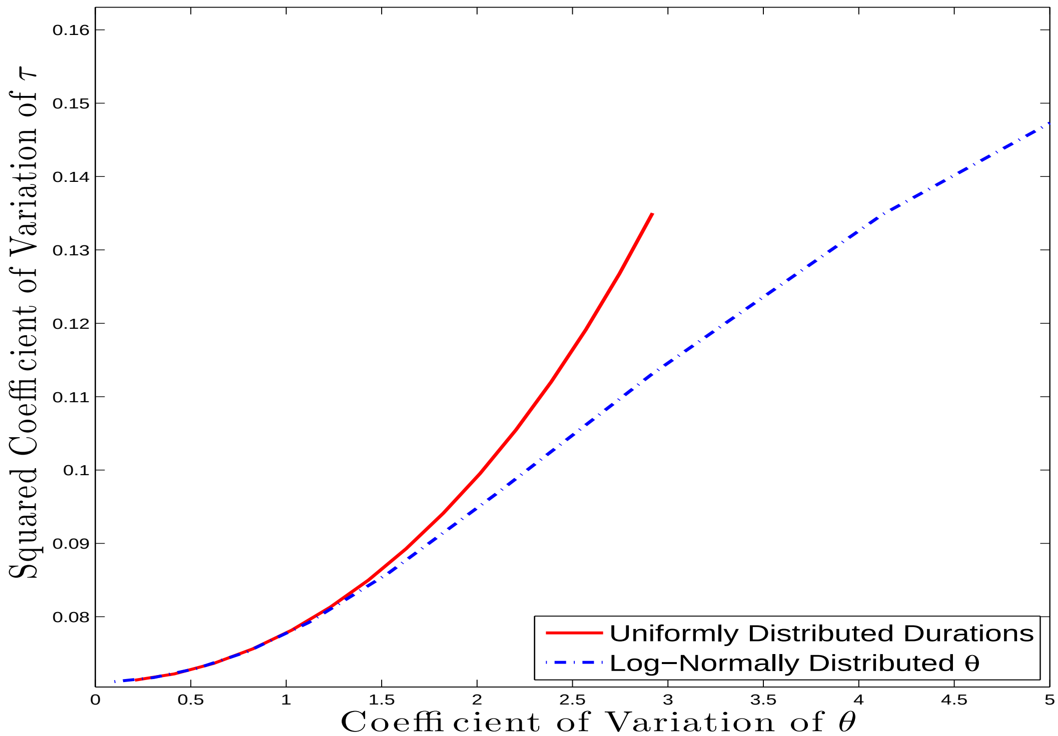
4.1 Variability of review times and observation costs

Proposition 7 characterizes the real effect of monetary shocks in terms of the variability of the times between reviews τ. In this section we use the characterization of the decision rules to express the variability of τ in terms of the variability of the observation cost θ. We use the baseline parametrization discussed in Section 3 and compare two alternative assumptions about the distribution of observation costs, $w(\theta)$, with respect to the case of an exponential distribution H studied in Section 3.2.1. In all cases considered we keep the average frequency of price reviews/adjustments per year constant at $N_a = 1.3$. In the first model we use equation (23) to compute the distribution $w(\theta)$ that produces an optimal times elapsed between observations $\hat{\tau}$ that is uniformly distributed. The uniform distribution is appealing because it only depends on two parameters, the minimum and maximum time between observations, and because it implies that different durations of reviews are equally likely, reflecting an agnostic prior. In the second model we assume that the distribution of θ is log-normal. Such a case is interesting because, differently from the previous case and from the exponential case, it does not allow for negative values of θ. In addition this distribution is flexible enough to allow for variability in the coefficients of variation of costs.

Figure 3 displays the coefficient of variation of θ that yields a given (squared) coefficient of variation of τ (on the vertical axis). The red solid line corresponds to the model in which $\hat{\tau}$ is uniformly distributed, while the dashed blue line corresponds to the model where θ is log-normal.[23] The figure shows that the variability of τ is increasing with the variability of θ even though, as suggested by the approximation in equation (17), the sensitivity of the coefficient of variation of τ with respect to the coefficient of variation of θ is small. In the model with uniformly distributed $\hat{\tau}$, there is an upper bound on the coefficient of variation of τ which comes from the natural restriction that $\tau \geq 0$. At this upper bound the coefficient of variation of θ is $CV(\theta) \approx 3$, and the implied cumulated output response, $\mathcal{M}(\delta)$, is only 14% larger than the value that a model with deterministic price reviews/adjustments would predict (e.g. Taylor), and about 43% smaller than it would be predicted by the case of exponential distributed reviews/adjustments (e.g. Mankiw-Reis or Calvo). In the model with a log-normally distributed θ the sensitivity of $CV(\tau)$ with respect to $CV(\theta)$ is even smaller than in the case of the uniformly distributed $\hat{\tau}$. For instance, when $CV(\theta) = 3$ then the implied cumulated output response, $\mathcal{M}(\delta)$, is only about 11% larger than the value

[23] Recall that τ is the actual duration of a review, while $\hat{\tau}$ is the duration of a review conditional on the product not being substituted. Given the exogenous arrival of substitutions, the first and second moments of τ needed to compute the coefficient of variation of τ are given by $\mathbb{E}(\tau) = \int_0^{E_v} \mathbb{E}(\tau|\ \theta)\ w(\theta)\ d\theta$ and $\mathbb{E}(\tau^2) = \int_0^{E_v} \mathbb{E}(\tau^2|\ \theta)\ w(\theta)\ d\theta$, where $w(\cdot) \equiv\ \phi((\log(\theta) - \mu_\theta)/\sigma_\theta)/(\theta\ \sigma_\theta \sqrt{2\pi})$ is the pdf of a log-normal, with $\phi(\cdot)$ being the pdf of a standard normal, $\mathbb{E}(\tau|\ \theta) = \int_0^{\hat{\tau}(\theta)} \tau\lambda e^{-\lambda\tau} + e^{-\lambda\hat{\tau}(\theta)}\hat{\tau}(\theta)$ and $\mathbb{E}(\tau^2|\ \theta) = \int_0^{\hat{\tau}(\theta)} \tau^2\lambda e^{-\lambda\tau} + e^{-\lambda\hat{\tau}(\theta)}(\hat{\tau}(\theta))^2$ are the moments conditional on a given realization of θ.

Figure 3: Variation of τ generated by transitory variation of θ



Note: The red solid line refers to the model where $\hat{H}(\tau)$ is uniform. The blue dashed line refers to the model with log-normal θ. In all cases parameters are such that $\mathbb{E}(\tau) = 1/1.3$. The other parameters are $\eta = 4$, $\sigma = 0.1$, $\gamma = \sigma^2/2$, $\lambda = 0.25$, $\rho = 0.02$ and $\mu = 0.00$.

that a model with deterministic reviews/adjustments would predict, and about 45% smaller than it would be predicted by the case of exponential distributed reviews/adjustments. If we increase the coefficient of variation of θ to $CV(\theta) \approx 23$, the squared coefficient of variation of τ only increases to about 1/4, where $\mathcal{M}(\delta)$ is only 25% larger than in the deterministic reviews/adjustments model (e.g. Taylor). At this parametrization, however, the average cost of a review is about 3.8% of profits, while the standard deviation is very high, about 87% of profits.[24]

In the case of a uniformly distributed $\hat{\tau}$ we find that an increase in the coefficient of variation for τ comes at cost of allowing for a significant mass of negative costs, $\theta < 0$, as was the case for the exponentially distributed $\hat{\tau}$. For instance, we find that producing a coefficient of variation in τ larger than $CV(\tau) = 1/4$ necessarily requires negative observation costs. Note moreover that that Proposition 6 shows that the excess kurtosis of price changes equals $3\,CV(\tau)^2$. Thus Figure 3 shows that to obtain large values of excess kurtosis (as the ones estimated in micro data sets) it is also required to have very large transitory fluctuations in observations costs. We interpret these results as evidence that large a variability in θ, allowing for negative values of the observation cost, in order to generate large variability in times between reviews.

5 Revisiting Reis' (2006) analysis on aggregation

As anticipated in Section 3, a key step in analyzing the predictions of costly information acquisition for the aggregate dynamics of the economy is to understand the implications of the firm-level policy rules for the economy-wide (or aggregate), response to shocks. To this end Propositions 6 and 7 in Reis (2006) present two different arguments stating that, under quite general assumptions on the form of individual-firm decision rules (H in our notation), the "distribution of inattentiveness" (Q in our notation) is exponential.

Appendix C shows that Reis's Proposition 6 is not correct. We claim that the correct result is Proposition 4 in Section 3.1. While according to Reis's claim any distribution H (the firm's distribution of reviews) can produce an exponential distribution of inattentiveness, we show that the conditions to obtain an exponential distribution of inattentiveness, Q, are very restrictive: only an H that is exponential itself will deliver an exponential Q, as Corollary 1 shows.

As an alternative foundation, Reis's (2006) Proposition 7 uses the Palm-Khintchine the-

[24] For comparison we note that in the case of an exponential $H(\tau)$ the coefficient of variation of the reviews is 1. Using the same parameter values we obtain that the associated coefficient of variation in θ supporting the exponential distribution of τ is $CV(\theta) \approx 11$.

orem to claim that the "distribution of inattentiveness" is exponential. In Appendix D we argue that the interpretation of this theorem for the problem under consideration is improper (while the theorem is of course correct). In particular, a correct interpretation of the theorem does not imply that the distribution of inattentiveness is exponential. Rather, we claim that imposing the assumptions of the Palm-Khintchine theorem in our setup leads to have almost all firms not changing prices during any finite time period.

The consequences of these findings are quite important. The fact that different form of randomness in the durations of reviews give rise to different distribution of inattentiveness, and therefore different aggregate dynamics in response to aggregate shocks, means that exact form of randomness can matter quite a lot. In addition, the importance for aggregate dynamics of the extent of randomness in the durations of reviews opens the door to the analysis of the foundations of such randomness, which we do in this paper.

6 Heterogenous time-invariant observation costs

Motivated by the existing literature, the previous sections studied a model where the cross section distribution of the times to the next review, Q, depended on the distribution of transitory i.i.d. variation in observation costs. In this section we study the other extreme case, namely a model where the cross section distribution of the times to the next review, Q, depends on permanent cross-firms variation in observation costs. Thus in this version of the model each firm has a time-invariant cost of observation $\tilde{\theta}$ which generates an optimal time to the next review $\hat{\tau}(\tilde{\theta})$. This is a special case of the firm's problem in equation (7) when the distribution of θ is degenerate at $\tilde{\theta}$. Solving for the optimal time to the next review, $\hat{\tau}$, we obtain one (implicit) equation for $\hat{\tau}$:

$$b\,\hat{\tau}(\tilde{\theta}) = \log\left(\frac{\left(1-e^{-(\rho+\lambda-b)\hat{\tau}(\tilde{\theta})}\right)/\,(\rho+\lambda-b)}{\left(1-e^{-(\rho+\lambda)\hat{\tau}(\tilde{\theta})}\right)/(\rho+\lambda)-\tilde{\theta}}\right). \tag{34}$$

where, to simplify the notation, we used $b \equiv \eta(\eta-1)\sigma^2/2$ and the usual assumption $\rho+\lambda > b$ applies. This is the same equation as in Reis's (2006) Proposition 5, although the notation is different. As Reis showed, permanent changes in $\tilde{\theta}$ produce changes in $\hat{\tau}$ with an elasticity of 1/2, at least for small $\tilde{\theta}$. In addition, when $\tilde{\theta} = 0$, then $\hat{\tau}(0) = 0$ -see Appendix B for more details. We notice two important differences with the case of temporary variation in observation cost analyzed in Proposition 3, and in particular with equation (14). First, $\hat{\tau}$ is concave in $\tilde{\theta}$ in the case of permanent variation, while it is convex in θ in the case of transitory variation. Second, there is no option value in the case of permanent variation, so that $\hat{\tau}(0) = 0$ according to equation (34). This implies that such model can generate arbitrarily small time between reviews with arbitrarily small, but still positive, observation cost.

Abusing notation, let $w(\tilde{\theta})$ be the cross section distribution of observation cost, so that $\int_0^x w(\tilde{\theta}) d\tilde{\theta}$ is the fraction of firms with permanent observation cost smaller than x. Using equation (34), there is a unique mapping from $w(\tilde{\theta})$ to the distribution of times between reviews, $\hat{H}$, which is given by

$$w(\tilde{\theta}) = -\hat{H}'\left(\hat{\tau}(\tilde{\theta})\right) \frac{\partial \hat{\tau}(\tilde{\theta})}{\partial \tilde{\theta}} \text{ for all } \theta > \underline{\theta} \, . \tag{35}$$

The distribution $\hat{H}(\tau)$ gives the *fraction of firms* that would choose to wait at least τ periods before the next review. Given the different nature of variability of observation cost, the distribution $\hat{H}(\tau)$ is a different object than in the case of i.i.d. transitory variation of observation costs. In the latter the function $\hat{H}(\tau)$ gave the *fraction of times* that any firm would choose to wait at least τ periods before the next review. Finally, let $Q(r)$ be the cross section distribution of times r until the next observation, as defined in Section 3, and $H(\tau)$ the cross section distribution of times to the next review including substitutions, $H(\tau) = \hat{H}(\tau)\, e^{-\lambda\tau}$. The mapping from the distribution H to the distribution Q is given, here too, by equation (18).

We next characterize the mapping from the coefficient of variation of costs $\tilde{\theta}$ to the coefficient of variation of times to the next review $\hat{\tau}$. Notice that here both the variation of costs and review times are taken with respect to the cross-section of firms. In fact, by assumption, there is no variation over time in either observation cost or time to the next review here. In Figure 4 we use equation (35) and compare three different assumptions about the distribution of observation costs, $w(\tilde{\theta})$: i) a distribution of observation costs $w(\tilde{\theta})$ producing an exponential distribution of times to the next review $\hat{H}$; ii) a distribution $w(\tilde{\theta})$ producing a uniform distribution $\hat{H}$; iii) log-normal $w(\tilde{\theta})$. In all cases, the parameters of the distribution $w(\tilde{\theta})$ are chosen so that the cross firms average frequency of price reviews/adjustments per year is $N_a = 1.3$. We plot the squared coefficient of variation of review times (including substitutions) against the coefficient of variation of observation costs.[25] A coefficient of permanent cross firms variation $CV(\tilde{\theta}) \approx 1.8$ is sufficient to support an exponential distribution of inattentiveness Q as assumed in Mankiw and Reis (2002). At our parameters, this is achieved with the average cost of a review being about 3.3% of profits, and the standard deviation being about 6% of profits. It is interesting to compare the results of this exercise to the corresponding exercise in the case of i.i.d. transitory variation in observation costs of

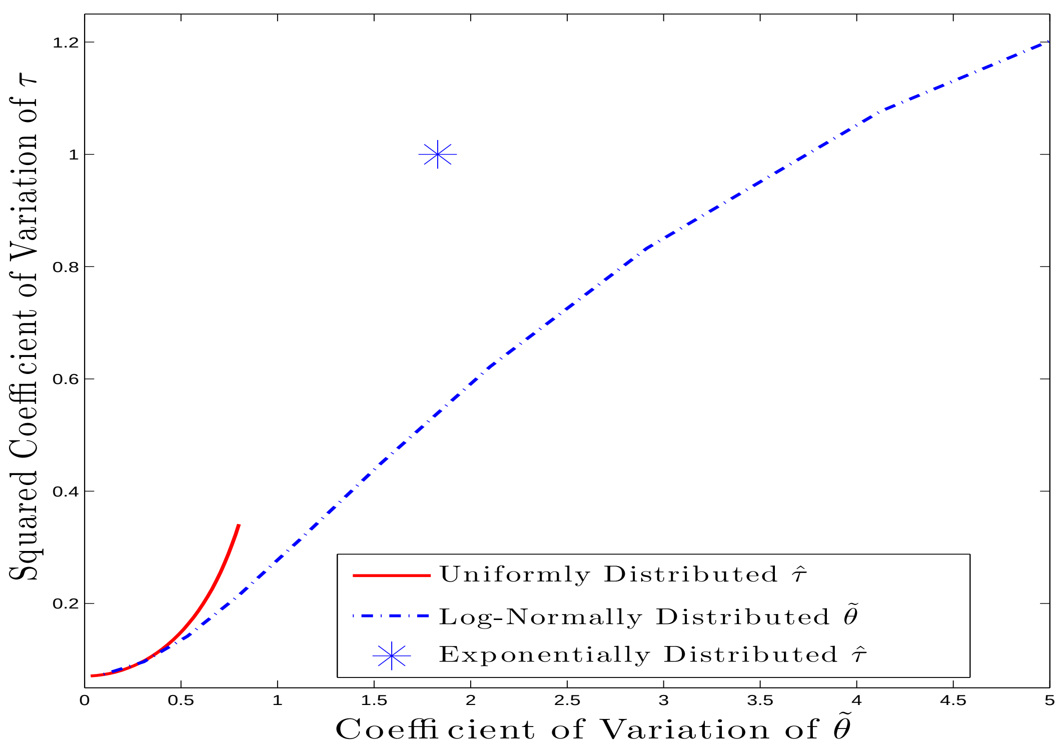
[25]The relationship between H and $\hat{H}$ is given by .

Section 4.1. The *cross section* sensitivity of $CV(\hat{\tau})$ with respect to permanent cross section variability in observation costs $CV(\tilde{\theta})$ is substantially larger than the sensitivity of *over time* variability in $\hat{\tau}$ with respect to i.i.d. transitory variability in observation cost, approximately given by by equation (17). This finding is analytically supported by Appendix B where we show that, given equation (34), the coefficient of (cross-sectional) variation in times to review is given by:

$$CV(\hat{\tau}) = \frac{1}{2} CV(\tilde{\theta}) + o(\bar{\theta}) \ , \tag{36}$$

where $\bar{\theta} \equiv \mathbb{E}(\tilde{\theta})$ is the cross-firms average cost of observation.

Figure 4: Variation of τ generated by permament variation of $\tilde{\theta}$



Note: The red solid line refers to the model where $\hat{H}(\tau)$ is uniform. The blue dashed line refers to the model where $\tilde{\theta}$ is log-normal. The star displays the coefficient of variation across $\tilde{\theta}$'s that supports an exponential $\hat{H}(\tau)$. In all cases parameters are such that $\mathbb{E}(\tau) = 1/1.3$. The other parameters of the model are $\eta = 4$, $\sigma = 0.1$, $\gamma = \sigma^2/2$, $\lambda = 0.25$, $\rho = 0.02$ and $\mu = 0.00$.

Finally we discuss implications of permanent variation in observation costs for impulse responses and distribution of price changes. First notice that the mapping from the distribution H to the cumulated output response, $\mathcal{M}(\delta)$, in equation (32) also apply here, despite the interpretation and definition of H are different as we discuss above.[26] Thus $\mathcal{M}(\delta)$ depends on the cross firms average frequency of observations/adjustments as well as on the cross firms coefficient of variation in times to the next review. Second, the cross sectional distribution of price changes is given by the same equations relating H to such distribution given in Section 3, but here the function H is a different object. In fact, as in the case of i.i.d. transitory variation in observation costs, cross firms heterogeneity in $\tilde{\theta}$ implies that the distribution of price changes across the whole economy is a mixture of normals, and hence is leptokurtic. Notice however that, differently from the case of i.i.d. transitory variation in observation costs, in this version of the model the distribution of price changes of a given firm is different from the cross section distribution of price changes. In fact, price changes for a given firm are normally distributed with variance $\sigma^2 \hat{\tau}(\tilde{\theta})$ and hence have kurtosis equal to 3.[27]

7 Concluding remarks

We studied a simple general equilibrium model of price setting where firms choose the time of price reviews facing transitory idiosyncratic variation in observation costs. We derived the mapping from the distribution of observation costs, $w(\theta)$, to the distribution of optimally chosen time elapsed between consecutive reviews, $H(\tau)$. Next, we derived the mapping from the distribution $H(\tau)$ for an individual firm to the invariant distribution of the times r until the next review for a cross section of firms, $Q(r)$, what Reis (2006) labels the "distribution of inattentiveness". Finally we derived a simple relationship between the distribution of inattentiveness and the cumulative output response to a monetary shock, and between the distribution of inattentiveness and the cross-sectional distribution of price changes.

We found that the real effects of monetary policy are smallest in the case of a constant observation cost, i.e. when the distribution of the times to review is degenerate. In this case the output response to a monetary shock is linear, as in a Taylor's (1980) model. We showed that the real effects of monetary policy increase if the observation costs are not constant: more variable times to review yields larger real effects of monetary policy. However the elasticity of the variance of the times to review with respect to the variability of the observation costs is very small, so that for moderate values of the volatility of the observation

[26]We note that the implications of permanent heterogeneity in the time between reviews are different from the ones produced by the heterogeneity in the Calvo parameter, which can be found in the literature (e.g. Carvalho (2006)). Here it is, essentially, heterogeneity in a deterministic length between adjustment, i.e. heterogeneity in the Taylor's length of contracts similar to Dixon and Kara (2011).

[27]Note also that in studies using scanner price data for a given firm the estimated kurtosis is higher than 3. For instance Midrigan (2011) finds a product level kurtosis equal to 4.

costs the macroeconomic effects are quantitatively close to the ones produced by a model with a constant cost. We show that to obtain large real effects of monetary shocks, such as those produced by the assumption of exponentially distributed adjustments, the model requires large variation of the observation costs, which must also include a substantial mass on *negative* observation costs. To better understand the effect of transitory shocks we compared the workings of this model with one with permanent shocks, i.e. a model with time invariant heterogeneity in the cross-section of review costs.

We characterized how the distribution of price changes depends on the distribution of the times to review. We found that the excess kurtosis of price changes is three times the squared coefficient of variation of the review times. This result, together with the characterization of the cumulated output response to a monetary shock, implies that if the distribution of times to review is chosen so that the kurtosis of price changes is consistent with the micro data, the effect of money on output is quite large, similar to the one implied by the exponential adjustments considered by Mankiw and Reis (2002).

To summarize, on the one hand we find that in this class of models the variation of observation cost required to produce important real effects of monetary shocks is quite large, we think implausibly so. On the other hand, we show that in this class of models the high kurtosis of the price changes found in the data is consistent with large effects of monetary shocks. We conclude that for these models to credibly imply large real effects of a monetary shock a mechanism different from variation in observation costs should be specified to produce large variation in times to review/adjust. Alternatively, if due to measurement error –of the type documented by Eichenbaum et al. (2012) for small price changes– the kurtosis of price changes were much smaller than what has been measured, then these models would imply small real effects of monetary shocks and would be consistent with plausible variation in observation cost. We leave the investigations of these alternatives for future work.

References

Alvarez, Fernando E., Herve Le Bihan, and Francesco Lippi. 2012. "Sticky prices, menu costs and the effect of monetary shocks: theory and micro-evidence." in preparation, EIEF.

Alvarez, Fernando E. and Francesco Lippi. 2012. "Price setting with menu cost for multi-product firms." NBER Working Papers 17923, National Bureau of Economic Research, Inc.

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

———. 2012. “Monetary Shocks with Observation and Menu Costs.” Tech. rep., University of Chicago mimeo.

Bhattarai, Saroj and Raphael Schoenle. 2010. “Multiproduct Firms and Price-Setting: Theory and Evidence from U.S. Producer Prices.” Working papers, Princeton University.

Bonomo, Marco and Carlos Carvalho. 2004. “Endogenous Time-Dependent Rules and Inflation Inertia.” *Journal of Money, Credit and Banking* 36 (6):1015–41.

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

Burstein, Ariel T. 2006. “Inflation and output dynamics with state-dependent pricing decisions.” *Journal of Monetary Economics* 53 (7):1235–1257.

Caballero, Ricardo J. and Eduardo M.R.A. Engel. 2007. “Price stickiness in Ss models: New interpretations of old results.” *Journal of Monetary Economics* 54 (Supplemen):100–121.

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

Carvalho, Carlos. 2006. “Heterogeneity in Price Stickiness and the Real Effects of Monetary Shocks.” *The B.E. Journal of Macroeconomics* 1.

Carvalho, Carlos and Felipe Schwartzman. 2012. “Selection and Monetary Non-Neutrality in Time-Dependent Pricing Models.” Tech. rep., FRB of Richmond.

Chari, V. V., Patrick J. Kehoe, and Ellen R. McGrattan. 2000. “Sticky Price Models of the Business Cycle: Can the Contract Multiplier Solve the Persistence Problem?” *Econometrica* 68 (5):pp. 1151–1179.

Cox, David R. and H.D. Miller. 2001. *The theory of stochastic processes*. Chapman and Hall, CRC, Boca Raton (first edition 1965).

Dixon, Huw and Engin Kara. 2011. “Contract length heterogeneity and the persistence of monetary shocks in a dynamic generalized Taylor economy.” *European Economic Review* 55 (2):280–292.

Eichenbaum, Martin S., Nir Jaimovich, Sergio Rebelo, and Josephine Smith. 2012. “How Frequent Are Small Price Changes?” NBER Working Paper 17956, National Bureau of Economic Research.

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

Heyman, D.P. and M.J. Sobel. 2003. *Stochastic models in operations research, Vol. I: Stochastic Processes and Operating Characteristics*, vol. 1. Dover Publications.

Klenow, J. Peter and Benjamin Malin. 2010. *Microeconomic Evidence on Price-Setting*. Handbook of Monetary Economics. Elsevier.

Mackowiak, B. and M. Wiederholt. 2009. "Optimal sticky prices under rational inattention." *American Economic Review* 99 (3):769–803.

Mackowiak, Bartosz Adam and Mirko Wiederholt. 2011. "Business cycle dynamics under rational inattention." Working Paper Series 1331, European Central Bank.

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.

———. 2010. *Imperfect Information and Aggregate Supply*, *Handbook of Monetary Economics*, vol. 3. Elsevier.

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

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

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

Sims, Christopher A. 2003. "Implications of rational inattention." *Journal of Monetary Economics* 50 (3):665–690.

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

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

A Proofs

Proof. (of Proposition 1)

1. The value of the policy $\tau = 0$ is $v(\theta) = -\theta + \int_{\underline{\theta}}^{\infty} v(\tilde{\theta})\, w(\tilde{\theta})\, d\tilde{\theta}$. Integrating both sides of this equation w.r.t. θ gives $\int_{\underline{\theta}}^{\infty} \theta\, w(\theta)\, d\theta = 0$.

2. If this was not the case, so that $v(\theta_1) < v(\theta_2)$ for $\theta_1 < \theta_2$, the firm could use the policy corresponding to the θ_2 at θ_1 and increase its value.

3. Suppose not, so that $\hat{\tau}(\underline{\theta}) = 0$, then $v(\underline{\theta}) = \int_{\underline{\theta}}^{\infty} v(\theta)\; w(\theta)\, d\theta - \underline{\theta}$, a contradiction with v being decreasing.

4. Let $\bar{\tau}$ and $\bar{g}$ be the optimal policy for the problem:

$$\bar{v} = \max_{\tau, g} \int_0^{\tau} e^{-(\rho+\lambda)s} f(g,s)\, ds \,+\, e^{-(\rho+\lambda)\tau + (\eta-1)(\gamma + (\eta-1)\frac{\sigma^2}{2})\tau} \left[\bar{v} - E_{\theta}\right] \;,$$

   The policy of setting $\hat{\tau}(\theta) = \bar{\tau}$ and $\hat{g}(\theta) = \bar{g}$ is feasible and has the same period cost when evaluated at $\theta = E_{\theta}$.

□

Proof. of Proposition 3

1. The foc for τ gives:

$$0 \geq e^{-(\rho+\lambda)\tau} \left(1 + e^{(\eta-1)\eta\frac{\sigma^2}{2}\tau} \left[-(\rho+\lambda) + (\eta-1)\eta\frac{\sigma^2}{2}\right] \left[E_v - \theta\right]\right) \quad \text{with} = \ \text{if } \tau > 0 \;. \tag{37}$$

2. The (unique) solution of the foc gives:

$$\hat{\tau}(\theta) = \begin{cases} \frac{\log\left(\left[(\rho+\lambda)-(\eta-1)\eta\frac{\sigma^2}{2}\right]\left[E_v - \theta\right]\right)}{-(\eta-1)\eta\frac{\sigma^2}{2}} > 0 & \text{if } \theta \leq E_v \\ \infty & \text{if } \theta > E_v \end{cases} \tag{38}$$

3. If $\underline{\theta} = 0$ then $\hat{\tau}(0) > 0$. Note that $E_v < 1/(\rho + \lambda - (\eta-1)\eta\frac{\sigma^2}{2})$ since the right hand side is the value of the case of not observation cost. Hence $\hat{\tau}(0) > 0$, as we argued for the general case.

4. Let θ_0 be the value of θ such that $\lim_{\theta \to \theta_0} \hat{\tau}(\theta) = 0$. From equation (38) we obtain

$$\theta_0 = E_v - 1/(\rho + \lambda - (\eta-1)\eta\frac{\sigma^2}{2}) < 0.$$

5. We start by writing the Bellman equation and replacing the FOC for $\hat{\tau}$ in it:

$$\begin{aligned} v(\theta) &= \frac{1-e^{-(\lambda+\rho)\hat{\tau}(\theta)}}{\lambda+\rho} + e^{-(\rho+\lambda)\hat{\tau}(\theta)} e^{(\eta-1)\eta\frac{\sigma^2}{2}\hat{\tau}(\theta)}[E_v-\theta], \\ &= \frac{1-e^{-(\lambda+\rho)\hat{\tau}(\theta)}}{\lambda+\rho} + \frac{e^{-(\rho+\lambda)\hat{\tau}(\theta)}}{(\rho+\lambda)-(\eta-1)\eta\frac{\sigma^2}{2}}, \\ &= \frac{1}{\rho+\lambda} + \left(\frac{(\eta-1)\eta\frac{\sigma^2}{2}}{(\rho+\lambda)(\rho+\lambda-(\eta-1)\eta\frac{\sigma^2}{2})}\right) e^{-(\lambda+\rho)\hat{\tau}(\theta)}, \end{aligned}$$

Taking expectations on both sides:

$$\mathbb{E}\left[v(\theta)\right] = \frac{1}{\rho+\lambda} + \left(\frac{(\eta-1)\eta\frac{\sigma^2}{2}}{(\rho+\lambda)(\rho+\lambda-(\eta-1)\eta\frac{\sigma^2}{2})}\right) \mathbb{E}\left[e^{-(\lambda+\rho)\hat{\tau}(\theta)}\right],$$

and using that

$$e^{-(\lambda+\rho)\hat{\tau}(\theta)} = \left(\left[(\rho+\lambda)-(\eta-1)\eta\frac{\sigma^2}{2}\right]\max\left(E_v-\theta,0\right)\right)^{\frac{\rho+\lambda}{(\eta-1)\eta\frac{\sigma^2}{2}}},$$

we obtain

$$\mathbb{E}\left[e^{-(\lambda+\rho)\hat{\tau}(\theta)}\right] = \left[(\rho+\lambda)-(\eta-1)\eta\frac{\sigma^2}{2}\right]^{\frac{\rho+\lambda}{(\eta-1)\eta\frac{\sigma^2}{2}}} \mathbb{E}\left[\max\left(E_v-\theta,0\right)^{\frac{\rho+\lambda}{(\eta-1)\eta\frac{\sigma^2}{2}}}\right].$$

Replacing back in the expected value of the value function:

$$E_v = \frac{1}{\rho+\lambda} + \\ \frac{(\eta-1)\eta\frac{\sigma^2}{2}}{(\rho+\lambda)(\rho+\lambda-(\eta-1)\eta\frac{\sigma^2}{2})}\left[(\rho+\lambda)-(\eta-1)\eta\frac{\sigma^2}{2}\right]^{\frac{\rho+\lambda}{(\eta-1)\eta\frac{\sigma^2}{2}}} \mathbb{E}\left[\max\left(E_v-\theta,0\right)^{\frac{\rho+\lambda}{(\eta-1)\eta\frac{\sigma^2}{2}}}\right],$$

we can write the expected value as:

$$\mathbb{E}\left[\max\left(E_v-\theta,0\right)^{\frac{\rho+\lambda}{(\eta-1)\eta\frac{\sigma^2}{2}}}\right] = \int_{\underline{\theta}}^{E_v} \left[E_v-\theta\right]^{\frac{\rho+\lambda}{(\eta-1)\eta\frac{\sigma^2}{2}}} w(\theta)d\theta,$$

obtaining one equation in one unknown, E_v:

$$E_v = \frac{1}{\rho+\lambda} + \frac{(\eta-1)\eta\frac{\sigma^2}{2}}{\rho+\lambda}\left[(\rho+\lambda)-(\eta-1)\eta\frac{\sigma^2}{2}\right]^{\frac{\rho+\lambda}{(\eta-1)\eta\frac{\sigma^2}{2}}-1}\int_{\underline{\theta}}^{E_v}(E_v-\theta)^{\frac{\rho+\lambda}{(\eta-1)\eta\frac{\sigma^2}{2}}}\,w(\theta)\,d\theta$$

Note that the rhs of this equation is strictly positive at zero. The first and second derivatives of the rhs of this expression w.r.t E_v are

$$\left[(\rho+\lambda)-(\eta-1)\eta\frac{\sigma^2}{2}\right]^{\frac{\rho+\lambda}{(\eta-1)\eta\frac{\sigma^2}{2}}-1}\int_{\underline{\theta}}^{E_v}(E_v-\theta)^{\frac{\rho+\lambda}{(\eta-1)\eta\frac{\sigma^2}{2}}-1}\,w(\theta)\,d\theta > 0,$$

$$\left[(\rho+\lambda)-(\eta-1)\eta\frac{\sigma^2}{2}\right]^{\frac{\rho+\lambda}{(\eta-1)\eta\frac{\sigma^2}{2}}-1}\left(\frac{\rho+\lambda}{(\eta-1)\eta\frac{\sigma^2}{2}}-1\right)\int_{\underline{\theta}}^{E_v}(E_v-\theta)^{\frac{\rho+\lambda}{(\eta-1)\eta\frac{\sigma^2}{2}}-2}\,w(\theta)\,d\theta > 0.$$

Note that the first derivative is zero at zero, it is strictly increasing, and if $E_v < 1/\left[\rho+\lambda-(\eta-1)\eta\frac{\sigma^2}{2}\right]$ it is strictly smaller than one. Hence the first intersection of the lhs and rhs occurs at a point wehre this inequality is satisfied. If there is a second intersection, it must be at a point where it is not satisfied, since the slope must be larger than one. Furthermore, there are no other intersections where the slope is smaller than one. To see that there is at most one intersection, divide the lhs and rhs by E_v raised to the power $\frac{\rho+\lambda}{(\eta-1)\eta\frac{\sigma^2}{2}} > 1$ and take E_v to infinity. The lhs goes go zero but the rhs goes to

$$\left[(\rho+\lambda)-(\eta-1)\eta\frac{\sigma^2}{2}\right]^{\frac{\rho+\lambda}{(\eta-1)\eta\frac{\sigma^2}{2}}-1}\lim_{x\to\infty}\int_{\underline{\theta}}^{x}\left(1-\frac{\theta}{x}\right)^{\frac{\rho+\lambda}{(\eta-1)\eta\frac{\sigma^2}{2}}}\,w(\theta)\,d\theta > 0.$$

Since the lhs and rhs must intersect at least once. Summarizing, there is at most one value where the lhs and rhs are equal and for which $E_v < 1/\left[\rho+\lambda-(\eta-1)\eta\frac{\sigma^2}{2}\right]$.

6. The solution of the foc is a maximum. Consider a case where $\theta < E_v$, then the second derivative of the right hand side of the value function,evaluated at $\hat{\tau}$ gives:

$$\begin{aligned} & e^{-(\rho+\lambda)\hat{\tau}}\left(-(\rho+\lambda)+e^{(\eta-1)\eta\frac{\sigma^2}{2}\hat{\tau}}\left[-(\rho+\lambda)+(\eta-1)\eta\frac{\sigma^2}{2}\right]^2[E_v-\theta]\right) \\ = \; & e^{-(\rho+\lambda)\hat{\tau}}\left(-(\rho+\lambda)-\left[-(\rho+\lambda)+(\eta-1)\eta\frac{\sigma^2}{2}\right]\right) = -e^{-(\rho+\lambda)\hat{\tau}}(\eta-1)\eta\frac{\sigma^2}{2} < 0. \end{aligned}$$

7. The monotonicity of $\hat{\tau}$ follows immediately from the effect of E_v in the optimal policy

in equation (14) and from the effect of w in E_v shown in Proposition 2.

8. Since $\mathbb{E}\left[\max\left(E_v - \theta, 0\right)^{\frac{\rho+\lambda}{(\eta-1)\eta\frac{\sigma^2}{2}}}\right]$ is the expected value of a convex function of θ for a fixed value of E_v, a mean preserving spread in $w(\cdot)$ increases its value. Hence a mean preserving spread in the distribution of θ increases the rhs of equation (15) for each E_v, and thus increases the value of the intersection.

□

Proof. (of Proposition 4)

The proof consists of considering a sequence of discrete time models with time period of length $x > 0$ and let x converge to zero. We use 'tildes' to denotes the objects of the discrete model with periods of length $x > 0$. Fix $x > 0$ and let $\tilde{H}(i; x) \equiv H(i\,x)$ for non-negative integers $i = 0, 1, 2, \cdots$ be the right CDF of the times between consecutive reviews. Likewise let $\tilde{R}(x)$ be the number of reviews per model period and let $\tilde{Q}(j\,;\,x)$ be the invariant distribution of times until the next review. The invariant distribution satisfies:

$$\tilde{Q}(j\,;\,x) = \tilde{R}(x) \sum_{i=0}^{\infty} \tilde{H}(j + \,i\,;\,x) \tag{39}$$

for $j = 0, 1, 2, \cdots$. We refer to the integers i and j as the number of model periods. We extend $\tilde{Q}(y; x)$, in the left-hand side of equation (39), to non-integer values of y by using the function $H(\cdot)$ i.e. $\tilde{H}(y + i; x) \equiv H((y + i)\,x)$. For integer values of j equation (39) can be understood as follows: consider the firms who reviewed their price i model periods ago, of which of there are a total of $\hat{R}(x)$. A fraction $\tilde{H}(j + i\,;\,x)$ of those firms have drawn times from which the next review will occur j model periods from now or later. Summing over the i index gives the total number of firms who will wait at least j periods before the next review. Note that $\tilde{Q}(0\,;\,x) = 1$ since, by assumption iii), review lengths are strictly higher than 0 with probability one. Evaluating equation (39) at $j = 0$, and using $\tilde{Q}(0; x) = 1$ we have:

$$1 = \tilde{Q}(0; x) = \tilde{R}(x) \sum_{i=0}^{\infty} \tilde{H}(i\,;\,x) \tag{40}$$

which gives one equation for $\tilde{R}(x)$:

$$\tilde{R}(x) = \frac{1}{\sum_{i=0}^{\infty} \tilde{H}(i\,;\,x)} \tag{41}$$

Defining by $R(x)$ the number of firms that review prices *per unit of time*, and by R its limit value, we have: $R(x) \equiv \frac{\tilde{R}(x)}{x}$ and $R \equiv \lim_{x \downarrow 0} R(x)$. Taking the limit in equation (41) as we shrink x

$$R = \lim_{x \downarrow 0} R(x) = \frac{1}{\lim_{x \downarrow 0} \sum_{i=0}^{\infty} H(i\,x)\,x} = \frac{1}{\int_0^{\infty} H(s)\,ds} \tag{42}$$

where we use the definition of $\tilde{H}$ and its relation to H, and the definition of the Riemann-Stieltjes integral. We define Q for any r as the limit $Q(r) = \lim_{x \downarrow 0} \tilde{Q}(r/x;\ x)$. Using equation (39), fixing a time $r \geq 0$ and defining x as $j = r\,x$ for any integer j we obtain:

$$\begin{aligned} Q(r) &= \lim_{x \downarrow 0} \frac{\tilde{R}(x)}{x} \sum_{i=0}^{\infty} H\left((j+i)x\right) x = \lim_{x \downarrow 0} R(x) \sum_{i=0}^{\infty} H\left(r + ix\right) x \\ &= R \int_0^{\infty} H\left(r+s\right) ds = R \int_r^{\infty} H\left(s\right) ds \end{aligned} \tag{43}$$

where j is integer, the first two equalities use the definition of x and the definition of $\tilde{H}$ and its relation to H, and the third uses the definition of a Riemann-Stieltjes integral. Differentiating equation (43), using the definition $q(r) = -Q'(r)$, and the expression for R we have

$$q(r) \equiv -Q'(r) = \frac{-\int_0^{\infty} H'(r+s)\,ds}{\int_0^{\infty} H(s)\,ds} = \frac{H(r)}{\int_0^{\infty} H(s)\,ds} \ . \tag{44}$$

□

Proof. (of Proposition 6) Consider first the distribution of log price changes of a firm reviewing/adjusting τ periods after the last observation/adjustment, $\Delta log(\ p(\tau)\ )$. Conditional on the next observation taking place in τ periods, the distribution of log-price changes upon the next observation/adjustment is normal with mean and variance equal to $(\mu - \gamma)\ \tau$ and $\sigma^2\ \tau$ respectively. Let index the distribution of $\Delta log(\ p(\tau)\ )$ by τ. Recall that the kurtosis of a random variable x with zero mean is defined as $Kurt(x) = m_4(x)/(m_2(x))^2$, where $m_s(x) = \mathbb{E}[x - \mathbb{E}(x)]^s$ is the s^{th} centered moment of x. Therefore, the second and fourth centered moments of $\Delta log(p(\tau))$ are equal to $m_2(\tau) = \sigma^2\ \tau$ and $m_4(\tau) = 3\ \sigma^4\ \tau^2$ respectively. Consider now the cross-sectional distribution of price changes, $\Delta \log(p)$. Such distribution is given by the mixture of normals arising from the different firms drawing different times to the next observation, τ, each with density $h(\tau)$.

The second moment of such mixture of normals is:

$$\int m_2(\tau) h(\tau) d\tau = \sigma^2 \int \tau\ h(\tau) d\tau.$$

which immediately gives equation (25). The fourth moment of such mixture of normals is given by the weighted average of the fourth moments of each normal, $m_4(\tau)$:

$$\int m_4(\tau)h(\tau)d\tau = 3\,\sigma^4 \int \tau^2 h(\tau)d\tau.$$

Thus the kurtosis of the distribution of log-price changes is given by:

$$Kurt(\Delta log(p)) = \frac{3\left(\int \tau^2 h(\tau)d\tau\right)}{\left(\int \tau h(\tau)d\tau\right)^2} = \frac{3\ (Var(\tau) + (1/N_a)^2)}{(1/N_a)^2} = 3\ (CV(\tau)^2 + 1)\,.$$

□

Proof. (of Proposition 5)

$$\begin{aligned}
w(\theta) &= \frac{\xi \exp\left(-\xi\,\hat{\tau}(\theta)\right)}{(\eta-1)\eta\frac{\sigma^2}{2}[E_v-\theta]} \\
&= \frac{\xi \exp\left(\frac{\xi}{(\eta-1)\eta\frac{\sigma^2}{2}}\ \log\left(\left[(\rho+\lambda)-(\eta-1)\eta\frac{\sigma^2}{2}\right][E_v-\theta]\right)\right)}{(\eta-1)\eta\frac{\sigma^2}{2}[E_v-\theta]} \\
&= \frac{\xi}{(\eta-1)\eta\frac{\sigma^2}{2}}\frac{\exp\left(\frac{\xi}{(\eta-1)\eta\frac{\sigma^2}{2}}\ \log\left(\left[(\rho+\lambda)-(\eta-1)\eta\frac{\sigma^2}{2}\right][E_v-\theta]\right)\right)}{[E_v-\theta]} \\
&= a_0\frac{\exp\left(\log\left(a_1^{a_0}[E_v-\theta]^{a_0}\right)\right)}{[E_v-\theta]} = a_0\frac{a_1^{a_0}\,[E_v-\theta]^{a_0}}{[E_v-\theta]} = a_0\,a_1^{a_0}\,[E_v-\theta]^{a_0-1}
\end{aligned}$$

where $a_0 = \frac{\xi}{(\eta-1)\eta\frac{\sigma^2}{2}}$. □

Proof. (of Proposition 7) First we show that

$$\int_0^\infty \int_t^\infty H(s)\,ds\,dt = \int_0^\infty s\,H(s)\,ds$$

We write $A(J) \equiv \int_0^J \int_t^J H(s)\,ds\,dt$ and $B(J) = \int_0^J s\,H(s)\,ds$. Note $A(0) = B(0) = 0$. Second, $B'(J) = JH(J)$ and $A'(J) = \int_0^J H(J)\,dt = JH(J)$, hence $A(J) = B(J)$ and taking $J \to \infty$ we obtained the desired result. Second let the variance of τ be

$$Var(\tau) \equiv -\int_0^\infty s^2 H'(s)ds - \left(-\int_0^\infty s\ H'(s)ds\right)^2 = 2\int_0^\infty s\ H(s)ds - \left(\int_0^\infty H(s)ds\right)^2,$$

where the second equality comes from integrating by parts. By using the first equation and $\int_0^\infty H(s)ds = 1/N_a$, the last equation immediately implies that we can express the cumulated output response in terms of variance and average time between reviews:

$$\mathcal{M}(\delta) = \frac{\delta}{\epsilon}\frac{1}{2}\left(1/N_a + \frac{Var(\tau)}{1/N_a}\right) \ . \tag{45}$$

□

B Coefficient of variation of optimal times to review

This appendix provides analytical approximations to characterize the coefficient of variation of optimal times to review in two limiting cases: first the case of a degenerate constant cost, $\theta = \tilde{\theta}$, in which changes in $\tilde{\theta}$ are permanent. In this case, the variation of times to review is intended with respect to the variability of the optimal choice across different firms. Second we consider temporary changes in θ, in the limiting case in which the variance becomes of θ becomes small. In this case, the variation of times to review is intended with respect to the variability of the optimal choice over time for a given firm.

B.1 The case of permanent variation in $\tilde{\theta}$

We show that the solution for $\hat{\tau}$ of equation (34) giving the optimal policy for the case of zero variance on the distribution of cost satisfies $\frac{\partial \log \hat{\tau}(0)}{\partial \log \tilde{\theta}} = \frac{1}{2}$. A solution of equation (34) gives the optimal policy $\hat{\tau}(\tilde{\theta})$. Exponentiating both sides of equation (34) and rearranging:

$$e^{-b\hat{\tau}}\left(1 - e^{-\hat{\tau}(\kappa-b)}\right) = (\kappa - b)\left(\frac{1 - e^{-\kappa\hat{\tau}}}{\kappa} - \tilde{\theta}\right) \tag{46}$$

where we use $b \equiv \eta(\eta-1)\sigma^2/2$ and $\kappa \equiv \rho + \lambda$. Expanding the exponentials in equation (46):

$$(1 - b\hat{\tau} - \cdots)\left(\hat{\tau}(\kappa-b) - \hat{\tau}^2\frac{(\kappa-b)^2}{2} + \cdots\right) = (\kappa - b)\left(\hat{\tau} - \hat{\tau}^2\frac{\kappa}{2} - \cdots - \tilde{\theta}\right)$$

Clearly, $\hat{\tau} = 0$ solves this equation for $\tilde{\theta} = 0$. In the case of $\tilde{\theta} > 0$. Dividing both sides by $\kappa - b$:

$$(1 - b\hat{\tau} + \cdots)\left(\hat{\tau} - \hat{\tau}^2\frac{(\kappa-b)}{2} + \cdots\right) = \left(\hat{\tau} - \hat{\tau}^2\frac{\kappa}{2} - \cdots - \tilde{\theta}\right)$$

ignoring terms of order higher than 2 in $\hat{\tau}$:

$$\hat{\tau} + \left(-b - \frac{(\kappa-b)}{2} + \frac{\kappa}{2}\right)\hat{\tau}^2 + o(\hat{\tau}^2) = \hat{\tau} - \tilde{\theta} + o(\hat{\tau}^2) \ ,$$

canceling the $\hat{\tau}$ terms in each side, and rearranging terms we obtain

$$\hat{\tau}(\tilde{\theta}) = \sqrt{\frac{2\,\tilde{\theta}}{\eta(\eta-1)\sigma^2/2} + o(\tilde{\theta})}\ . \tag{47}$$

Let $\bar{\theta} \equiv \mathbb{E}(\tilde{\theta})$. Assuming a constant strictly positive coefficient of variation $\sqrt{\nu} > 0$, so that $Var(\tilde{\theta}) = \nu\,\bar{\theta}^2$, and using equation (47), $\mathbb{E}(\hat{\tau}) = \hat{\tau}(\bar{\theta}) + o(\bar{\theta})$ and $Var(\hat{\tau}) = \left(\hat{\tau}'(\bar{\theta})\right)^2 Var(\theta) + o\left(Var(\theta)\right)$ we obtain

$$\begin{aligned}\frac{Var(\hat{\tau})}{(\mathbb{E}(\hat{\tau}))^2} &= \frac{1}{2\bar{\theta}/b + o\left(\bar{\theta}\right)}\left(\left(\frac{2\bar{\theta}}{b} + o\left(\bar{\theta}\right)\right)^{-1/2}\frac{1}{b}\right)^2 Var(\tilde{\theta}) + \frac{o\left(\nu\,\bar{\theta}^2\right)}{2\bar{\theta}/b + o\left(\bar{\theta}\right)} \\ &= \frac{1}{4}\,\frac{Var(\tilde{\theta})}{(\bar{\theta})^2} + o\left(\bar{\theta}\right).\end{aligned}$$

B.2 The case of temporary variation in θ

We characterize how the variance of transitory changes in observation costs affects the variance of $\hat{\tau}$. Let us first derive an expression for the derivative of $\hat{\tau}(\theta)$, evaluated at the mean of the observation cost $\bar{\theta} \equiv \mathbb{E}(\theta)$. Direct computation using equation (14) gives $\hat{\tau}'(\bar{\theta}) = \frac{1}{b(v-\bar{\theta})}$ where we use that in the limiting case of a zero variance $E_v = v$. Integrating the value function in equation (13) for the case of zero variance gives

$$v - \bar{\theta} = \frac{\left(1 - e^{-\kappa\hat{\tau}}\right)/\kappa - \bar{\theta}}{1 - e^{-(\kappa+b)\hat{\tau}}}$$

Combining the right hand side of this equation with equation (34) allows us to write $v - \bar{\theta} = \frac{e^{-b\hat{\tau}}}{\kappa - b}$ which then gives

$$\hat{\tau}'(\bar{\theta}) = \frac{\kappa - b}{b}\,e^{b\hat{\tau}} = \frac{\kappa - b}{b}\,e^{\sqrt{2\bar{\theta}b}} \tag{48}$$

where the last equality uses the square root formula derived above for the limit case of no variance. Next, assume a constant strictly positive coefficient of variation $\sqrt{\nu} > 0$, so that $Var(\theta) = \nu\,\bar{\theta}^2$. When $\nu \approx 0$, $\hat{\tau}(\theta)$ is approximately given by equation (47) so that $(\mathbb{E}(\hat{\tau}))^2 = 2\bar{\theta}/b + o\left(\bar{\theta}\right)$. Using equation (48), and $Var(\hat{\tau}) = \left(\hat{\tau}'(\bar{\theta})\right)^2 Var(\theta) + o\left(Var(\theta)\right)$ we obtain

$$
\begin{aligned}
\frac{Var(\hat{\tau})}{(\mathbb{E}(\hat{\tau}))^2} &= \left(\frac{\kappa-b}{b}\right)^2 \left(1+\sqrt{2\bar{\theta}b+o\left(\bar{\theta}\right)}+\frac{1}{2!}\left(2\bar{\theta}b+o\left(\bar{\theta}\right)\right)+...\right)^2 \frac{Var(\theta)}{2\bar{\theta}/b+o\left(\bar{\theta}\right)}+\frac{o\left(\nu\,\bar{\theta}^2\right)}{2\bar{\theta}/b+o\left(\bar{\theta}\right)} \\
&= \left(\frac{\kappa-b}{b}\right)^2 \left(1+o\left(\sqrt{\bar{\theta}}\right)\right)\frac{Var(\theta)}{2\bar{\theta}/b\left(1+\frac{o\left(\bar{\theta}\right)}{2\bar{\theta}/b}\right)}+o\left(\bar{\theta}\right) \\
&= \left(\frac{\kappa-b}{b}\right)^2 \frac{Var(\theta)}{\bar{\theta}^2}\frac{b\bar{\theta}}{2}\left(1+o\left(\sqrt{\bar{\theta}}\right)\right)\left(1-\frac{o\left(\bar{\theta}\right)}{\bar{\theta}}\right)+o\left(\bar{\theta}\right) \\
&= (\kappa-b)^2\,\nu\frac{\bar{\theta}}{2b}\left(1+o\left(\sqrt{\bar{\theta}}\right)-\frac{o\left(\bar{\theta}\right)}{\bar{\theta}}-\frac{o\left(\bar{\theta}^{\frac{3}{2}}\right)}{\bar{\theta}}\right)+o\left(\bar{\theta}\right) \\
&= (\kappa-b)^2\,\nu\frac{\bar{\theta}}{2b}+o\left(\bar{\theta}\right)
\end{aligned}
$$

which, using $\nu=\frac{Var(\theta)}{(\bar{\theta})^2}>0$, gives equation (17).

C A correction of Reis' Proposition 6

Proposition 6 in Section 5 of Reis (2006) states that if the process $f_i(t)$ that describes the time elapsed between observations of an individual firm i is the same one for all firms, and it is independent across time and across producers, then the "distribution of inattentiveness" is exponential. We claim that the proposition is incorrect and show that the "distribution of inattentiveness" is exponential if and only if the firm level process $f_i(t)$ is exponential. A simple counter-example suggests that Proposition 6 cannot be correct: suppose that the time between reviews for all firms has a bounded support, so that for instance no firm would ever draw a review time longer than a year, then the invariant distribution cannot converge to exponential, which has an unbounded support. We reproduce a time invariant version of Reis's (2006) Section 5 set-up (albeit with different notation), obtain the relevant expressions for the invariant distribution, and use it to construct some counterexamples to proposition 6. Our conclusion is that one must *assume* Poisson distributed reviews, as opposed to deriving it from any variation on the length of reviews, as it is claimed in Reis's (2006) *Proposition 6*.

We end this section by first reviewing the notation used in Reis (2006), and secondly by comparing his results (Proposition 6) to ours (Proposition 4 and Corollary 1). Regarding the notation, Reis (2006) refers to the calendar times of consecutive i^{th} -1 and i^{th} reviews as $D(i-1)$ and $D(i)$, and to the length of the time elapsed between the two reviews as $d(i)=D(i)-D(i-1)$. Reis's (2006) Proposition 6 uses the assumption in Reis's (2006)

Assumption 2 which states that the time between consecutive reviews d are i.i.d. distributed, with density $f_i(t)$, and common and independent across all firms (page 805 : "The arrival of decision dates then takes the form of a stochastic point process. Its properties are described by a set of probability density functions for how long the inattentiveness period will last, conditional on when the producer last adjusted.") While Reis's (2006) Proposition 6 does not have a formal definition of "inattentiveness", the paragraph just after makes clear that it refers to the distribution of the age of reviews: "The process of arrival of decision dates is, therefore, a Poisson process with parameter ρ. That is, if at any point in time, we survey the producers on how long ago they last planned, we will find that the share not having planned for x periods equals $\rho e^{-\rho x}$..." Additionally, while the proof does not define "inattentiveness" formally, the first expression of step 5 shows that it aims to compute the distribution of V_τ which is previously defined as the remaining duration of the plan. As noted in Cox and Miller (2001), this is the same as the distribution of the age of the plans.

Importantly, our Corollary 1 stands in stark contradiction with Reis's (2006) Proposition 6, which states that the distribution of inattentiveness is exponential "without further assumptions" than the one described in Reis's (2006) Assumption 2. As our corollary shows, the distribution of inattentiveness is exponential if and only if the distribution of elapsed time between reviews is exponential.

D A comment on Reis' Proposition 7

In this section we reconsider the results in Proposition 7 of Reis (2006), which uses the Palm-Khintchine theorem. We argue that Reis's interpretation of the theorem is incorrect. Following Reis (2006), we consider an economy composed of J firms, each of them with a renewal process describing its reviews, and analyze the resulting aggregate renewal process for the whole economy. Proposition 7 of Reis (2006) considers the limit case of an economy with J firms having times between reviews that are i.i.d. for each firm, independent across firms, and distributed with a right CDF $H_{j,J}$. The expected value of the reciprocal of the time between reviews for firm j is denoted by $R_{j,J}$. The sub-index j indicates the firm, and while the sub-index J indicates the number of firms in the economy. It is useful to define two stochastic processes associated to a given distribution $H_{j,J}$. The first is $N_{j,J}(t)$, the number of events (reviews) that have occurred from time zero to time t. The second is the sequence of random variables $\tau_{j,J}$, denoting the length of time elapsed between consecutive events (reviews). The set-up is the one of the super-imposition of independent point process, leading to the Palm-Khintchine theorem.[28]

The objective is to characterize the distribution of times since the last review for the *whole* economy, as J goes to infinity. The superimposition of J independent renewal processes is defined as $N_J(t) \equiv N_{1,J}(t) + N_{2,J}(t) + \cdots + N_{J,J}(t)$. Correspondingly, τ_J is the time elapsed for the *whole* economy between consecutive reviews, undertaken by any of the J firms, and given by $\tau_J \equiv \min_{j=1,\ldots,J}\{\tau_{j,J}\}$. Its associated right CDF is given by $H_J(\tau) = \prod_{j=1}^{J} H_{j,J}(\tau)$ (see expression (83) in Cox and Miller (2001)). The Palm-Khintchine theorem characterizes the limit of the superposition of J independent renewal processes as $J \to \infty$, and the distribution of the corresponding times between events. Under an extra assumption, this theorem shows that as $J \to \infty$ the process $N_J(t)$ converges to a *Poisson* process with intensity $R = \lim_{J\to\infty} \sum_{j=1}^{J} R_{j,J} < \infty$.

We make two remarks about the application of this theorem. First, for R to be finite, the average frequency of reviews must go to zero for almost all firms. In other words, this is an economy where the time between consecutive reviews for the representative firms is, on average, arbitrarily long. The second comment uses the law of large numbers to interpret probabilities across firms as fractions. In particular, the fraction of firms conducting a review in a interval of time of length dt converges to zero, as $dt \to 0$ and $J \to \infty$. This is because the *number* of firms reviewing in an interval of length dt is approximately $R\,dt$.

A simple example which satisfies the assumptions of Reis's (2006) Proposition 7 of illustrates these two points. Suppose that each of the J processes are independent *Poisson* processes with intensity $R_{j,J} = R/J$. In this case one does not even need the Palm-Khintchine theorem, since it is well known (and easy to show) that the superposition (sum) of J Poisson processes is itself a Poisson process with arrival rate equal to the sum of the arrival rates, i.e. with rate R. As J becomes large this economy converges to one with no expected reviews for each firm, and with an infinite time between reviews. Note that for any J, the expected number of events (reviews) per unit of time in the whole economy is R, but since there are J firms, the expected number of reviews per firm is R/J, which goes to zero as $J \to \infty$.

From the analysis above, we conclude that the fraction of firms that review in any given period of time converges to zero under the assumption of Proposition 7 in Reis (2006). This result is used incorrectly in Reis's (2006) derivation of the Phillips curve in section 6.3. In fact , Proposition 7 is used as to imply that the distribution of the age of the reviews ($H(j)$ in the notation of pages 808 and 809) is exponential. As discussed above, under the assumptions of Proposition 7, the distribution of the age of reviews across all the firms is degenerate as the number of firms in the economy goes to infinity under the conditions of Proposition 7.

[28]See Heyman and Sobel (2003) section 5.8 starting in page 156 or Cox and Miller (2001) section 9.5.i starting in page 363.

E Correlated idiosyncratic demand and cost shocks

This appendix considers the problem of a firm that faces correlated idiosyncratic demand and cost shocks z, so that the demand level A_t is now given by $A_t \; z^{1-\eta}$. We introduce this problem for several reasons: in a setup where demand and cost shocks are correlated the relative demand of the different consumption goods can be made stationary, and in particular the frictionless profit can be made constant. In this case all the variation of $\theta_t(z) = \theta \, \Pi_t^*$ (as in equation (6)) is due to variation in θ. Furthermore the model with correlated demand and cost shocks has been analyzed in the literature by several authors (see Woodford (2009), Bonomo, Carvalho, and Garcia (2010) Midrigan (2011), Alvarez and Lippi (2012)), so that exploring this case serves as a robustness check of the results for our benchmark case.

The price setting equation is as before, i.e. equation (1), but equation (2) is replaced by

$$\Pi_t(p) \;\; = \;\; A_t \; F\left(p/p^*\right) \qquad \text{and} \qquad \Pi_t^* = A_t \; F(1) \tag{49}$$

where, letting $g \equiv \frac{p}{p^*}$, the function $F(g) \equiv \left(g \, \frac{\eta}{\eta-1}\right)^{-\eta} \left(g \, \frac{\eta}{\eta-1} - 1\right)$. As before from the definition of g, and from the laws of motion of W and z, that the dynamics of g are given by equation (4), i.e. $d\log(g(t)) = (\gamma - \mu) \; dt + \sigma dB(t)$ where $B(t)$ is a Wiener process. In steady state, i.e. $A_{t+s} = A_t$, expected static profits evaluated s periods after re-setting the price to a price gap g at time t, and scaled by the maximum static monopolist profits, Π^*, at time t, are denoted by $f(g, s)$:

$$f(g,s) \equiv \mathbb{E}\left[\frac{F(g_{t+s})}{F(1)} \; \Big| g_t = g\right] = \eta \, g^{1-\eta} \, e^{\left((\eta-1)(\mu-\gamma)+\frac{\sigma^2}{2}(\eta-1)^2\right)s} - (\eta-1) \, g^{-\eta} \, e^{\left(\eta(\mu-\gamma)+\frac{\sigma^2}{2}\eta^2\right)s} \; , \tag{50}$$

The first term of equation (50) depends on expected growth in real revenues between t and $t+s$. The second term of equation (50) depends on expected growth in real marginal cost between t and $t+s$. Notice that both the growth in revenues and marginal cost are stochastic as the shock z is both a shock to productivity and a shock to demand. In this case the value function for the firm is

$$v(\theta) \;\; = \;\; \max_{\tau \in \bar{\mathbb{R}}_+, g \geq 0} \int_0^\tau e^{-(\rho+\lambda)s} \; f(g,s) \, ds \; + e^{-(\rho+\lambda)\,\tau} \left[-\theta + \int_{\underline{\theta}}^\infty v(\tilde{\theta}) \, w(\tilde{\theta}) \, d\tilde{\theta}\right] \; . \tag{51}$$

Notice that, compared to equation (7), the model with the 2 shocks is simpler in that the terms that depend on the evolution of the level of the demand (through productivity) drop out of the equation (both directly and though the discount factor of the continuation).

It is easily shown that a first order approximation of equation (50) around $f(1,0)$ gives

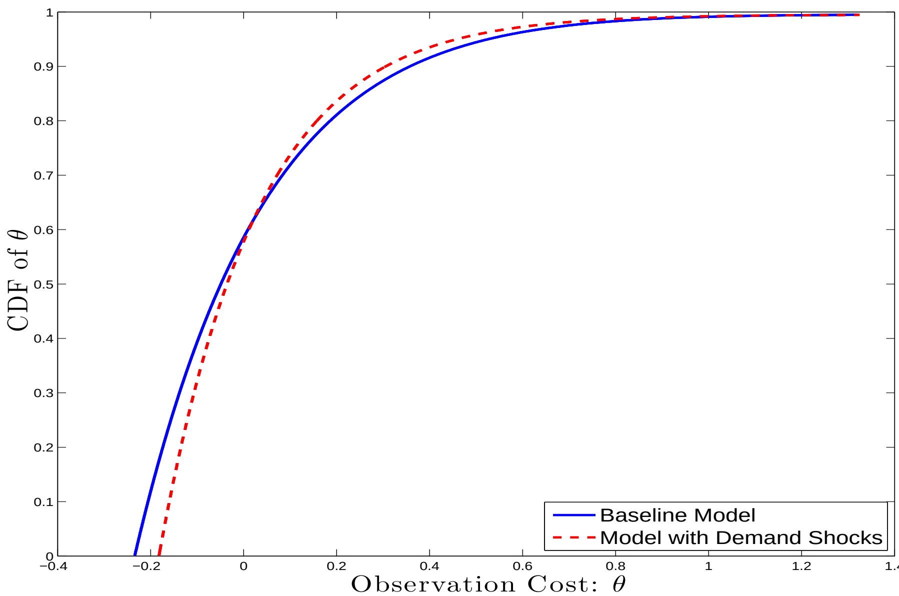
$$f(g,s) = 1 - \frac{\eta(\eta-1)}{2}\sigma^2 \; s + o(||\sigma^2 s, (g-1)||)$$

In this case the problem is similar to the one we studied in Alvarez, Lippi, and Paciello (2011), in the sense that the value function for the firm problem with zero inflation (equation 7 on page 1928 in our paper) is obtained if the function $f(g,s)$ is approximated by $\sigma^2 s$ which is obviously a good approximation for small values of t around the optimal return point $g = 1$.[29]

Next we present some numerical results for the model with 2 costs, and compare to our baseline model with idiosyncratic productivity only. In this analysis, we use the same parametrization of Figure 1 and Figure 3, so that $\rho = 0.02$, $\eta = 4$, $\sigma = 0.1$, $\lambda = 0.25$, $\mu = 0$ and $\gamma = \sigma^2/2$, and the average frequency observations/adjustments is of 1.3 per year. This parametrization produces a distribution of price changes with a standard deviation equal to 0.085. Figure 5 plots the CDF of θ, $\int_{\underline{\theta}}^{\theta} w(\tilde{\theta})\, d\tilde{\theta}$, that supports an exponential distribution of adjustment times $\hat{H}(\tau)$, in the model with two shocks against our baseline model with one shock. Qualitatively, the two distributions are very similar. Quantitatively, the two models put similar mass on negative observation costs. Figure 6 plots the square of the coefficient of variation of τ against the coefficient of variation of θ, in the case of log-normally distributed θ, in the model with two shocks against our baseline model with one shock. Again, the two models predict similar relationship between variation in τ and variation in θ, for given average time between reviews.

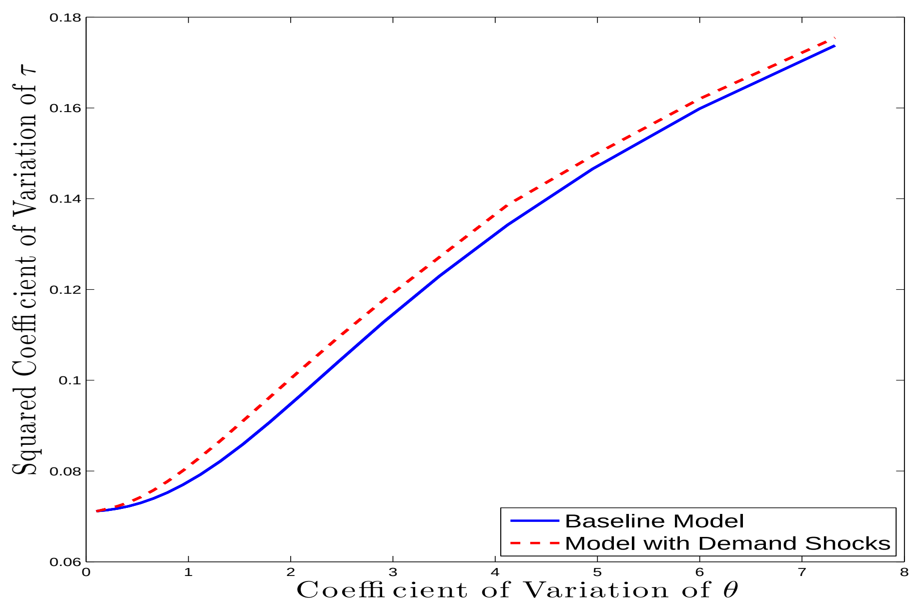
[29] For clarity of comparison note that the value function in Alvarez, Lippi, and Paciello (2011) is expressed as a difference from the maximized frictionless profits, i.e. $f(g,s) - 1$.

Figure 5: Distribution of θ supporting an exponential distribution of adjustment times $\hat{H}(\tau)$



Note: The figures plots the CDF of θ: $\int_{\underline{\theta}}^{\theta} w(\tilde{\theta})\ d\tilde{\theta}$. The red dashed line refers to the model with two idiosyncratic shocks, while the blue line refers to the baseline model with one idiosyncratic shock. The distribution of observation costs implies an invariant distribution of the age of reviews which is exponential with $\xi = 1.05$ observations/price changes per year. Since there are also $\lambda = 0.25$ substitutions per year, the baseline case has 1.3 price changes per year. Other parameters of the model are $\eta = 4$, $\sigma = 0.10$, $\gamma = \sigma^2/2$, $\lambda = 0.25$, $\rho = 0.02$.

Figure 6: Variation of τ against variation of θ



Note: Coefficient of variation is the standard deviation of a random variable divided by its mean. The red dashed line refers to the model with two idiosyncratic shocks, while the blue line refers to the baseline model with one idiosyncratic shock. The distribution of θ is log-normal, with each coefficient of variation in θ corresponding to a different parametrization of the log-normal. In all cases parameters are such that $\mathbb{E}(\tau) = 1/1.3$. The other parameters of the model are $\eta = 4$, $\sigma = 0.1$, $\gamma = \sigma^2/2$, $\lambda = 0.25$, $\rho = 0.02$ and $\mu = 0.00$.