

NBER WORKING PAPER SERIES

INFORMATION-CONSTRAINED STATE-DEPENDENT PRICING

Michael Woodford

Working Paper 14620

<http://www.nber.org/papers/w14620>

NATIONAL BUREAU OF ECONOMIC RESEARCH

1050 Massachusetts Avenue

Cambridge, MA 02138

December 2008

I would like to thank Marco Bonomo, Ariel Burstein, Ricardo Caballero, Eduardo Engel, Michael Golosov, Bob King, John Leahy, Bart Mackowiak, Filip Matejka, Giuseppe Moscarini, Emi Nakamura, Chris Sims, Tony Smith, Jon Steinsson and Alex Wolman for helpful discussions; Maxim Pinkovskiy and Luminita Stevens for outstanding research assistance; the NSF for research support through a grant to the NBER; and the Arthur Okun and Kumho Visiting Professorship, Yale University, for providing the time to begin work on this project. The views expressed herein are those of the author(s) and do not necessarily reflect the views of the National Bureau of Economic Research.

© 2008 by Michael Woodford. All rights reserved. Short sections of text, not to exceed two paragraphs, may be quoted without explicit permission provided that full credit, including © notice, is given to the source.

Information-Constrained State-Dependent Pricing
Michael Woodford
NBER Working Paper No. 14620
December 2008
JEL No. E31

ABSTRACT

I present a generalization of the standard (full-information) model of state-dependent pricing in which decisions about when to review a firm's existing price must be made on the basis of imprecise awareness of current market conditions. The imperfect information is endogenized using a variant of the theory of “rational inattention” proposed by Sims (1998, 2003, 2006). This results in a one-parameter family of models, indexed by the cost of information, which nests both the standard state-dependent pricing model and the Calvo model of price adjustment as limiting cases (corresponding to a zero information cost and an unboundedly large information cost respectively). For intermediate levels of the information cost, the model is equivalent to a “generalized Ss model” with a continuous “adjustment hazard” of the kind proposed by Caballero and Engel (1993a, 1993b), but provides an economic motivation for the hazard function and very specific predictions about its form. For high enough levels of the information cost, the Calvo model of price-setting is found to be a reasonable approximation to the exact equilibrium dynamics, except in the case of (infrequent) large shocks. When the model is calibrated to match the frequency and size distribution of price changes observed in microeconomic data sets, prices are found to be much less flexible than in a full-information state-dependent pricing model, and only about 20 percent more flexible than under a Calvo model with the same average frequency of price adjustment.

Michael Woodford
Department of Economics
Columbia University
420 W. 118th Street
New York, NY 10027
and NBER
michael.woodford@columbia.edu

Models of state-dependent pricing [SDP], in which not only the size of price changes but also their timing is modeled as a profit-maximizing decision on the part of firms, have been the subject of an extensive literature.¹ For the most part, the literature dealing with empirical models of inflation dynamics and the evaluation of alternative monetary policies have been based on models of a simpler sort, in which the size of price changes is modeled as an outcome of optimization, but the timing of price changes is taken as given, and hence neither explained nor assumed to be affected by policy. The popularity of models with exogenous timing [ET] for such purposes stems from their greater tractability, allowing greater realism and complexity on other dimensions. But there has always been general agreement that an analysis in which the timing of price changes is also endogenized would be superior in principle.

This raises an obvious question: how much is endogeneity of the timing of price changes likely to change the conclusions that one obtains about aggregate dynamics? Results available in special cases have suggested that it may matter a great deal. In a dramatic early result, Caplin and Spulber (1987) constructed a tractable example of aggregate dynamics under SDP in which nominal disturbances have no effect whatsoever on aggregate output, despite the fact that individual prices remain constant for substantial intervals of time. Danziger (1999) obtains a similarly stark neutrality result, again for a special case allowing a closed-form solution, but this time with idiosyncratic as well as aggregate shocks. The Caplin-Spulber and Danziger examples are obviously extremely special; but Golosov and Lucas (2007) find, in numerical analysis of an SDP model calibrated to account for various facts about the probability distribution of individual price changes in U.S. data, that the predicted aggregate real effects of nominal disturbances are quite small, relative to what one might expect based on the average interval of time between price changes. And more recently, Caballero and Engel (2007) consider the real effects of variation in aggregate nominal expenditure in a fairly general class of “generalized Ss models,” and show that quite generally, variation in the “extensive margin” of price adjustment (*i.e.*, variation in the number of prices that adjust, as opposed to variation in the amount by which each of these prices changes) implies a smaller real effect of nominal disturbances than would be predicted in an ET model (and hence variation only on the “intensive margin”); they argue that the degree of immediate adjustment of the overall level of

¹See, for example, Burstein and Hellwig (2007), Dotsey and King (2005), Gertler and Leahy (2007), Golosov and Lucas (2007), Midrigan (2008), and Nakamura and Steinsson (2008a, 2008b) for some recent additions.

prices can easily be several times as large as would be predicted by an ET model.²

These results suggest that it is of some urgency to incorporate variation in the extensive margin of price adjustment into models of the real effects of monetary policy, if one hopes to obtain results of any quantitative realism. Yet there is one respect in which one may doubt that the results of standard SDP models are themselves realistic. Such models commonly assume that at each point in time, each supplier has completely precise information about current demand and cost conditions relating to its product, and constantly re-calculates the currently optimal price and the precise gains that would be obtained by changing its price, in order to compare these to the “menu cost” that must be paid to actually change the price. Most of the time no price change is justified; but on the first occasion on which the benefit of changing price becomes as large as the menu cost, a price change will occur. Such an account assumes that it is *only* costs associated with actually changing one’s price that are economized on by firms that change prices only infrequently. Instead, studies such as Zbaracki *et al.* (2004) indicate that there are substantial costs associated with information gathering and decisionmaking that are also reduced by a policy of reviewing prices only infrequently.³ If this is true, the canonical SDP model (or “Ss model”), according to which a price adjustment occurs in any period if and only if a certain adjustment threshold has been reached, should not yield realistic conclusions. In fact, a model that takes account of the costs of gathering and processing information is likely to behave in at least some respects like ET models.⁴ The question is to what extent a more realistic model of this kind would yield conclusions about aggregate price adjustment and the real effects of nominal disturbances that are similar to those of ET models, similar to those of canonical SDP models, or different from both.

The present paper addresses this question by considering a model in which the

²An earlier draft of their paper (Caballero and Engel, 2006) proposed as a reasonable “benchmark” that the degree of flexibility of the aggregate price level should be expected to be about three times as great as would be predicted by an ET model calibrated to match the observed average frequency of price changes.

³Zbaracki *et al.* report that at the firm that they studied, the total managerial costs of reviewing the firm’s pricing policy are 7 times as large as the physical cost of changing the posted prices.

⁴Phelps (1990, pp. 61-63) suggests that ET models may be more realistic than SDP models on this ground. Caballero (1989) presents an early analysis of a way in which costs of information acquisition can justify “time-dependent” behavior, which is further developed by Bonomo and Carvalho (2004) and Reis (2006).

timing of price reviews is determined by optimization subject to an information constraint, in a dynamic extension of the model proposed in Woodford (2008). The model generalizes the canonical SDP model (which appears as a limiting case of the more general model, the case of zero information cost) to allow for costs of obtaining and/or processing more precise information about the current state of the economy, between the intermittent occasions on which full reviews of pricing policy are undertaken. For the sake of simplicity, and to increase the continuity of the present contribution with prior literature, it is assumed that when a firm decides to pay the discrete cost required for a full review of its pricing policy, it obtains *full information* about the economy’s state at that moment; hence when price changes occur, they are based on full information, as in canonical SDP models (as well as canonical ET models).⁵ However, *between* the occasions on which such reviews occur, the firm’s information about current economic conditions is assumed to be much fuzzier; and in particular, the decision *whether to conduct a full review* must be made on the basis of much less precise information than will be available after the review is conducted. As a consequence, prices do not necessarily adjust at precisely the moment at which they first become far enough out of line for the profit increase from a review of pricing policy to justify the cost of such a review.

There are obviously many ways in which one might assume that information is incomplete, each of which would yield somewhat different conclusions. Here (as in Woodford, 2008) I adopt a parsimonious specification based on the concept of “rational inattention” proposed by Sims (1998, 2003, 2006). It is assumed that *all* information about the state of the world is equally available to the decisionmaker — one does not assume that some facts are more easily or more precisely observable than others — but that there is a limit on the decisionmaker’s ability to process information of *any* kind, so that the decision is made on the basis of rather little information. The information that the decisionmaker obtains and uses in the decision is, however, assumed to be the information that is *most valuable* to her, given the decision problem that she faces, and subject to a constraint on the *overall rate of information flow* to

⁵The assumption that full information about current conditions can be obtained by paying a fixed cost also follows the previous contributions of Caballero (1989), Bonomo and Carvalho (2004), and Reis (2006); I depart from these authors in assuming that *partial* information about current conditions is also available between the occasions when the fixed cost is paid. The analysis here also differs from theirs in assuming that access to memory is costly, as discussed further in section 1.2.

the decisionmaker. This requires a quantitative measure of the information content of any given indicator that the decisionmaker may observe; the one that I use (following Sims) is based on the *information-theoretic* measure (entropy measure) proposed by Claude Shannon (1948).⁶ The degree of information constraint in the model is then indexed by a single parameter, the cost per unit of information (or alternatively, the shadow price associated with the constraint on the rate of information flow). I can consider the optimal scheduling of price reviews under tighter and looser information constraints, obtaining both a canonical SDP model and a canonical ET model as limiting cases; but the more general model treated here introduces only a single additional free parameter (the information cost) relative to a canonical SDP model, allowing relatively sharp predictions.

The generalization of the canonical SDP model obtained here has many similarities with the “generalized Ss model” of pricing proposed by Caballero and Engel (1993a, 2007) and the SDP model with random menu costs of Dotsey, King and Wolman (1999). Caballero and Engel generalize a canonical Ss model of pricing by assuming that the probability of price change is a continuous function of the signed gap between the current log price and the current optimal log price (*i.e.*, the one that would maximize profits in the absence of any costs of price adjustment), and estimate the “adjustment hazard function” that best fits US inflation dynamics with few *a priori* assumptions about what the function may be like. The model of price-adjustment dynamics presented in sections 1 and 2 below is of exactly the form that they assume. However, the “hazard function” is given an economic interpretation here: the randomness of the decision whether to review one’s price in a given period is a property of the *optimal* information-constrained policy. Moreover, the model here makes quite specific predictions about the form of the optimal hazard function: given the specification of preferences, technology and the cost of a review of pricing policy, there is only a one-parameter family of possible optimal hazard functions, corresponding to alternative values of the information cost. For example, Caballero and Engel assume that the hazard function may or may not be symmetric and might equally well be asymmetric in either direction; this is treated as a matter to be de-

⁶See, *e.g.*, Cover and Thomas (2006) for further discussion. The appendix of Sims (1998) argues for the appropriateness of the Shannon entropy measure as a way of modeling limited attention. As is discussed further in section 1.2, the informational constraint assumed here differs from the one proposed by Sims in the way that memory is treated.

terminated empirically. In the model developed here, the hazard function is predicted to be asymmetric in a particular way, for any assumed value of the information cost.

Caballero and Engel (1999) propose a structural interpretation of generalized Ss adjustment dynamics (in the context of a model of discrete adjustment of firms' capital stocks), in which the cost of adjustment by any given firm is drawn independently (both across firms and over time) from a continuous distribution of possible costs; Dotsey, King and Wolman (1999) [DKW] consider the implications for aggregate price adjustment and the real effects of nominal disturbances of embedding random menu costs of this kind in a DSGE model with monopolistically competitive pricing. The predicted dynamics of price adjustment in the model developed here are essentially the same as in a particular case of the DKW model; there exists a particular distribution for the menu cost under which the DKW model would imply the same hazard function for price changes as is derived here from optimization subject to an information constraint.⁷

However, the present model supplies an alternative interpretation of the randomness of adjustment at the microeconomic level that some may find more appealing than the idea of random menu costs. Moreover, the present model makes much sharper predictions than the DKW model; there is only a very specific one-parameter family of menu-cost distributions under which the DKW model makes predictions consistent with the information-constrained model. Assumptions that appear completely arbitrary under the random-menu-cost interpretation (why is it natural to assume that the menu cost should be i.i.d.?) are here derived as a consequence of optimization. At the same time, assumptions that might appear natural under the random-menu-cost interpretation (a positive lower bound on menu costs, or a distribution with no atoms) can here be theoretically excluded: the optimal hazard function in this model necessarily corresponds to a distribution of menu costs with *an atom at zero*. This has important implications: contrary to the typical prediction of parametric versions of the Caballero-Engel or DKW model, the present model implies that there is always (except in the limit of zero information cost) a positive

⁷Like the DKW model, the present model implies in general that the adjustment hazard should be a monotonic function of the amount by which the firm can increase the value of its continuation problem by changing its price. Only in special cases will this allow one to express the hazard as a function of the signed gap between the current log price and the optimal log price, as in the "generalized Ss" framework of Caballero and Engel (1993a, 1993b). Section 2, however, offers an example of explicit microfoundations for such a case.

adjustment hazard even when a firm’s current price is *exactly optimal*. This makes the predicted dynamics of price adjustment under the present model more similar to those of the Calvo (1983) model than is true of these other well-known generalizations of the canonical SDP model. It also helps to explain the observation in microeconomic data sets of a large number of very small price changes, as stressed by Midrigan (2008),⁸ and increases the predicted real effects of nominal disturbances (for a given overall frequency of price change), for reasons explained by Caballero and Engel (2007).

In fact, the results obtained here suggest that the predictions of ET models may be more reliable, for many purposes, than results from the study of SDP models have often suggested. The Calvo (1983) model of staggered price-setting is derived as a limiting case of the present model (the limit of an unboundedly large information cost); hence this model, often regarded as analytically convenient but lacking in any appealing behavioral foundations, can be given a fully explicit decision-theoretic justification — the quantitative realism of which, relative to other possible specifications, then becomes an empirical matter. Moreover, even in the more realistic case of a positive but finite information cost, the model’s prediction about the effects of typical disturbances can be quite similar to those of the Calvo model, as is illustrated numerically below. The present model predicts that the Calvo model will be quite inaccurate in the case of *large* enough shocks — large shocks should trigger immediate adjustment by almost all firms, because even firms that allocate little attention to monitoring current market conditions between full-scale reviews of pricing policy should notice when something dramatic occurs — and in this respect it is surely more realistic than the simple Calvo model. Yet the shocks for which this correction is important may be so large as to occur only infrequently, in which case the predictions of the Calvo model can be quite accurate much of the time.

Section 1 characterizes the optimal timing of reviews of pricing policy in a stylized model with a structure similar to that assumed by Caballero and Engel (1993a, 1993b); this analysis extends the model of information-constrained discrete choice proposed in Woodford (2008) to an infinite-horizon dynamic setting. Section 2 then

⁸Midrigan (2008) proposes an alternative explanation to the one given here for a positive hazard function when the current price is nearly optimal (see Figure 4 of his paper). The present model achieves a similar effect, without the complication of assuming interdependence between price changes for different goods.

illustrates the application of this general framework to a specific model of monopolistically competitive price-setting with idiosyncratic shocks. Section 3 compares the numerical predictions of a calibrated version of this model to microeconomic evidence regarding individual price changes. Section 4 then discusses the implications of the model for the neutrality of money, and section 5 concludes.

1 Rational Inattention in a Dynamic Model of the Timing of Discrete Adjustments

Here I present a dynamic extension of the model of information-constrained discrete adjustment presented (in a one-period context) in Woodford (2008). As in the work of Caballero and Engel (1993a, 2007), I shall simplify the state space by considering a “tracking” problem, in which a firm’s profits each period depend only on its “normalized price,”⁹ *i.e.*, the difference between its log price p_t and the current value of a state variable m_t outside the firm’s control, about which it is only imperfectly informed. (For example, profits may depend on the firm’s price relative to its current unit cost of production. An explicit model of monopolistically competitive price adjustment with the structure assumed in this section is presented in section 2.)

In this section, I consider the scheduling of price reviews by a single firm. (An equilibrium with many firms simultaneously making similar decisions is treated in section 2.) The model is one with a countably infinite sequence of discrete dates (indexed by integers t) at which the firm’s price may be adjusted (and at which sales occur).¹⁰ I shall suppose that the firm seeks to maximize the expected value of a

⁹The definition given here of a stationary optimal policy can be extended in a relatively straightforward way to the case in which profits also depend on other variables, including aggregate state variables. But the notation is simplified in this presentation by abstracting from such additional state variables, and it allows us to obtain a model in which the adjustment hazard is a function solely of a “price gap,” as in the work of Caballero and Engel. It also considerably simplifies the numerical analysis in section 3, as is discussed further in section 4.

¹⁰The model could be extended in a reasonably straightforward way to the scheduling of reviews of pricing policy in continuous time, as in Reis (2006). But discrete time is mathematically simpler and allows more direct comparison with much of the prior literature on state-dependent pricing.

discounted objective function of the form

$$\sum_{t=0}^{\infty} \beta^t \pi(q_t), \quad (1.1)$$

where $q_t \equiv p_t - m_t$ is the firm's (log) normalized price in period t , and single-period profits are assumed to be given by a function $\pi(q)$ that reaches its unique maximum at an interior value that can be normalized as $q = 0$.

Uncertainty about the firm's normalized price results from the random evolution of the state m_t representing market conditions. Again in order to reduce the size of the state space required to characterize equilibrium dynamics (and again following Caballero and Engel), I shall assume for the sake of simplicity that this evolves according to an exogenously given random walk,

$$m_t = m_{t-1} - z_t, \quad (1.2)$$

where the innovation z_t is drawn independently each period from a probability distribution with density function $g(z)$. (The sign of the innovation is chosen so that a positive innovation z_t corresponds to an increase in q_t .)

1.1 Information Constraints

I shall suppose that the (log) price p_t charged by the firm reflects current and past information about the evolution of m_t in three distinct ways. First, I suppose that the firm *reviews its pricing policy* only at certain times, rather than constantly. Holding such a review involves a substantial fixed cost, which is the reason that reviews occur only as discrete events; but when a review is held, payment of the fixed cost allows the firm to collect a great deal of information about market conditions at that time, on the basis of which the new pricing policy is chosen. Second, between the occasions on which a review is conducted, the firm charges a price for its product in accordance with its current pricing policy. The information about current conditions that can be used in the implementation of such a policy — that is, the extent to which p_t can depend on the current state of the world, as opposed to instructions written down at the time of the last review — is assumed to be quite limited. (This is why it matters that policy is not more frequently updated.) And third, the decision about *whether to conduct a review* of pricing policy is made on the basis of incomplete information

about current conditions. How well the firm's price p_t will track variations in m_t depends in general on what one assumes about the amount of information used in the decision about the scheduling of price reviews, the amount of information obtained when conducting such reviews, and the amount of additional information that can be used in implementing the pricing policy chosen as a result of the review.

The focus of the present paper is the *price review* decision; hence the information used in that decision will be considered in detail, while I make extremely simple assumptions about the available information for the other two purposes. First, I shall assume that at the time of a price review, payment of the fixed cost gives the firm access to *full information* about the current state of the economy. Hence the new plan that is chosen is the optimal one under that state of the world, as is commonly assumed in models with exogenous timing of price reviews, whether these involve a fixed price between reviews (as in the models of Taylor or Calvo) or some more complex plan (as in the models of Mankiw and Reis, 2002, or Devereux and Yetman, 2003); as well as in models with state-dependent timing of reviews, again whether these involve a fixed price (as in standard menu-cost models) or a more complex plan (as in the model of Burstein, 2006); and in generalizations of state-dependent pricing of the kind proposed by Caballero and Engel (1993a, 2007) or Dotsey, King and Wolman (1999). This assumption not only simplifies the analysis of the consequences of a particular timing for the price reviews, but also allows me to obtain standard models of price adjustment (both a standard “Ss” model and the Calvo model) as limiting cases of the model considered here.

Second, I assume that the pricing policies that are implemented between reviews use *no information* about the current state of the economy; hence the pricing policy reduces to a *single price* that is charged until the policy is reviewed. (Technically, the zero-information assumption would allow the firm to choose to randomize each period over a set of prices, as long as the randomization is independent of the current state; but I assume a single-period profit function $\pi(q)$ such that it is never optimal to randomize.¹¹) Hence the dynamics of price adjustment in this model are the same as in a model in which one must pay a “menu cost” to change one's price, and the menu cost is also the fixed cost of obtaining new (complete) information about the state of the economy on the basis of which to set the new price. However, I prefer to interpret the model as one in which there are no true menu costs, but only several

¹¹Specifically, $\pi(q)$ is a strictly concave function of a monotonically increasing function of q .

types of information costs. For one thing, if one supposes that the price is fixed between reviews of pricing policy owing to a menu cost, there is no very plausible reason to suppose that the same fixed cost should both allow one to change one's price and to obtain information that one would otherwise not have. Moreover, it is fairly common in some retail sectors to observe pricing policies which do *not* involve a single price, but (for example) frequent alternations between two or more prices, even though the *set* of prices among which the seller alternates remains unchanged for many months; such behavior suggests a model in which (i) the pricing *policy* is reconsidered only at fairly long intervals, and in which (ii) the pricing policy involves only a very coarse discrimination among different weeks, so that only a few different prices are ever charged,¹² but the relative insensitivity of the price to changing market conditions reflects information costs (inattentiveness) rather than menu costs. While more complex pricing policies of that sort are not considered in this paper, a model that allowed for them would represent an interesting extension of the simpler theory developed here.

It is also important to note that I shall treat awareness of the passage of *time* as among the types of information about the current state of the world that may be costly for the firm. (Given that the information constraints are interpreted as limits on the *attention* of the decisionmaker, and not as claims about what it is difficult to observe in the world, the fact that it is easy to construct accurate time-keeping devices is irrelevant to this issue.) This means that when I consider the incomplete information on the basis of which the review scheduling decision is made, I assume that a firm *may* choose to know how long a time has elapsed since its last review, but that this information is costly in the same way as other sorts of information about its current circumstances. My assumption thus differs from that of Bonomo and Carvalho (2004) or Reis (2006), who assume that information about random events since the last review¹³ cannot be used in deciding whether to schedule a review,

¹²Matejka (2008) shows how an information-flow constraint can result in a policy that alternates among a small number of prices.

¹³What I am calling the dates of “reviews” correspond to the dates at which information is updated in the model of Reis (2006). Reis’s model, however, is equivalent to one in which pricing plans are chosen at discrete dates (the dates at which information is updated) and followed until the next information update; under one of these plans, the price charged each period may depend on the time that has elapsed since the last information update, but not on any random events that have occurred since information was last updated. The decision about when to update information again

but that this decision may depend on the length of time that has elapsed since the last review. Similarly, when I assume that a pricing policy must use no information about the current state, this means not only that the price charged cannot depend on unforecastable changes in conditions since the adoption of the policy, but also that the price cannot depend on the time that has elapsed. In this my assumption differs from both Reis (2006) and Burstein (2006), who allow firms to follow pricing policies under which the price is a deterministic function of time between reviews, though it may not depend on any other information about the current state.¹⁴

1.2 Rational Inattention

I turn now to the precise specification of the information used in the scheduling of reviews of pricing policy. I adopt Sims' (1998, 2003, 2006) hypothesis of "rational inattention": firms have precisely that information that is most valuable to them, given the decision problem that they face, subject to a constraint on the overall quantity of information that they access. Rather than specifying a quantity constraint, I assume that there is a cost $\theta > 0$ per unit of information obtained each period by the decisionmaker, and that the total quantity of information obtained is the amount that is optimal given this cost.

I assume that the firm can arrange to receive a (possibly multi-dimensional) signal each period, which may be related in a fairly arbitrary way to the state of the economy at that time. Let $\omega_t \in \Omega_t$ denote a complete description of the economy's state in period t (including the complete history of all disturbances to that date). The firm arranges to receive a signal s_t drawn from some set S , where the conditional probability $\pi(s_t|\omega_t)$ of receiving any given signal is chosen in advance by the firm; the firm's decision about whether to review its pricing policy in period t is then a (possibly random) function of the signal s_t that is received. The cost to the firm per period of receiving this signal is θI , where $I \geq 0$ is the Shannon (1948) measure of

is also allowed to depend on the time that has elapsed since the last update, but not on random events that have occurred since then.

¹⁴Note that Burstein's model, like the one proposed here (but unlike the model of Reis), is one in which it is assumed that pricing plans must use less information than is used in deciding whether it is time to revise one's plan. (In Burstein's model, the revision decision is state-dependent, while the plan that is adopted is not.) But I assume that information is more costly than in Burstein's model, in the case of *each* of these decisions.

the average information content of the signal, namely the average amount by which the entropy of the posterior distribution over Ω_t (after observing the signal s_t) is less than the entropy of the firm's prior. Here the entropy measure of the uncertainty indicated by a given probability density f over the state space Ω_t is defined as

$$-E[\log f(\omega_t)],$$

where the expectation is under the distribution f . The parameter θ indicates the degree to which the decisionmaker's attention is scarce, with a higher value of θ requiring the decisionmaker to economize on attention to a greater extent (and hence to use less information).

A first elementary result (Woodford, 2008) is that under an optimal information structure, the signal s_t will take only two possible values, and can be interpreted as a "yes/no" signal as to whether the current period is a good time to review one's pricing policy. Since the only use of the signal is to decide whether to conduct the review, a signal that differentiates more finely among states will convey redundant information; and since the more informative signal would have a greater cost (will place a greater burden on the decisionmaker's attention) without improving the quality of the decision, it would be inefficient. Similarly, an optimal price-review policy will necessarily be a deterministic function of the signal (*i.e.*, a price review is always conducted if and only if the signal is "yes"); for in the case that arbitrary randomization of the decision is desired, it is more efficient to arrange for this by increasing the randomness of the signal (lowering its information content and hence its cost), rather than by randomizing after receiving an unnecessarily informative assessment of market conditions. This means that an informationally efficient price-review strategy (where the design of the signalling mechanism is treated as part of this strategy) can be fully described by specifying a *hazard function* $\Lambda_t(\omega_t)$ that indicates the probability of a price review in any state.

Another elementary result (again see Woodford, 2008) is that an informationally efficient strategy will involve signals that convey information only about aspects of the state that are relevant to the firm's decision problem. In the present problem, this means that the probability of receiving a given signal $s \in S$ will depend only on the current value of q_t , and not on prior history or any other aspects of the current state. Hence the strategy that is followed in any period can be described by a hazard function $\Lambda_t(q_t)$, a measurable function of a single real variable taking values in $[0, 1]$,

as in the papers of Caballero and Engel (1993a, 2007).

The information cost of a strategy represented by a hazard function $\Lambda_t(q_t)$ depends on the firm's prior over possible values of q_t before the signal s_t is received. In a dynamic model, this depends on the price-review strategy followed by the firm in earlier periods. (If price reviews are very frequent, the firm should know that it is unlikely that its normalized price has drifted very far from the value to which it would have been reset if a price review has occurred in the recent past, while if they are infrequent, the firm should be much more uncertain about its current normalized price.) Hence it is necessary to model the consequences of a price-review strategy for the evolution of the prior.

Under the assumptions summarized above, the normalized price of firm i evolves according to

$$q_{t+1}(i) = q_t(i) + z_{t+1} \quad (1.3)$$

if there is no review of the firm's policy in period t , whereas

$$q_{t+1}(i) = q_t^* + z_{t+1}$$

if firm i reviews its policy in period t . Here $q_t(i)$ is the normalized price of firm i in period t , *after* realization of the period t change in m_t , but *before* the decision about whether to review the firm's policy in period t , and q_t^* is the normalized price (*after* the review) that is chosen by a firm that reviews its policy in period t . The value of q_t^* is the same for all firms i because (as is shown below) the optimal choice for a firm that reviews its policy is independent of the normalized price that it has at the time of the review; hence if firms differ only in the periods in which they happen to have reviewed their prices in the past (despite having followed identical policies), q_t^* will be the same for all i .

The dynamics of a firm's prior also depend on what we assume about the firm's memory. In the applications of rational inattention to dynamic decision problems by Sims (1998, 2003, 2006), memory of the entire history of past signals is assumed to be perfectly precise (and costless); the information-flow constraint applies only to the degree of informativeness of *new* observations of external reality. Instead, I shall assume that access to one's own memory is *as costly* as access to any other source of information, during the intervals between price reviews, and this includes memory of the time at which one last reviewed one's pricing policy. For example, one may allow firms to condition their price-review decision on the number of periods n that have

elapsed since the last price review. In this case, the firm has a prior $f(q, n)$ over the joint distribution of its current normalized price q and the current value of n for that firm. The firm can learn the value of n and condition its decision on that value, but this would have an information cost of

$$-\theta \sum_n f_n \log f_n,$$

where $f_n \equiv \int f(q, n) dq$ is the marginal prior distribution over values of n . Assuming that the unit cost θ of this kind of information is identical to the cost of information about the value of q , the firm will optimally choose not to learn the current value of n ; since learning the value of n would be of use to the firm *only* because this information would allow it to estimate the current value of q with greater precision, it would always be more efficient to use any information capacity that it devotes to this problem to observe the current value of q with greater precision, rather than bothering to observe the value of n .

In assuming that the cost of information about the firm's memory of its own past signals is exactly the same as the cost of information about conditions external to the firm, I am making an assumption that is fully in the spirit of Sims' theory of rational inattention: rather than assuming that some kinds of information are easily observable while others are hidden, the cost of *any* kind of information is assumed to be the same as any other, because the relevant bottleneck is limited *attention* on the part of the decisionmaker, rather than anything about the structure of the world that obscures the values of certain state variables. This is admittedly a special case, but it is the assumption that makes Sims' theory such a parsimonious one. It is also a convenient case to analyze first, owing to its simplicity.¹⁵

In this case, any firm i begins any period t with a prior $f_t(q)$ over the possible values of $q_t(i)$. This prior indicates the *ex ante* distribution of possible values of the

¹⁵Interestingly, the literature on informational complexity constraints in game theory has more often made the *opposite* choice to that of Sims: it is considered more natural to limit the information content of a decisionmaker's memory than the information content of her perception of her current environment. For example, in Rubinstein (1986) and many subsequent papers, it is assumed that a strategy (in a repeated game) is preferred if it can be implemented by a finite-state automaton with a smaller number of states; this means, if it requires the decisionmaker to discriminate among a smaller number of different possible histories of previous play. But while memory is in this sense assumed to be costly, there is assumed to be no similar advantage of a strategy that reduces the number of different possible observations of current play among which the decisionmaker must discriminate.

firm's normalized price in period t , given the policy followed in previous periods, but not conditioning on any of the signals received in previous periods, or on the timing of previous price reviews. By “the policy” followed in previous periods, I mean the design of the signalling mechanism, determining the probabilistic relation between the state and the signal received each period, and the firm's intended action in the event that any given signal is received, but not the history of the signals that were actually received or the actions that were taken. Given the discussion above, the policy followed in period t can be summarized by a hazard function $\Lambda_t(q)$, indicating the probability of a price review in period t as a function of the normalized price in that period, and a reset value q_t^* , indicating the normalized price that the firm chooses if it reviews its pricing policy in period t . As a result of this policy, $q_{t+1}(i)$, the normalized price in period $t + 1$ (after realization of the period $t + 1$ innovation in market conditions) will be equal to $q_t^* + z_{t+1}$ with probability $\Lambda(q_t(i))$ and equal to $q_t(i) + z_{t+1}$ with probability $1 - \Lambda(q_t(i))$, conditional on the value of $q_t(i)$. Integrating over the distribution of possible values of $q_t(i)$, one obtains a prior distribution for period $t + 1$ equal to

$$f_{t+1}(q) = g(q_t^* - q) \int \Lambda_t(\tilde{q}) f_t(\tilde{q}) d\tilde{q} + \int g(\tilde{q} - q) (1 - \Lambda_t(\tilde{q})) f_t(\tilde{q}) d\tilde{q}. \quad (1.4)$$

This is the prior at the beginning of period $t + 1$, regardless of the signal received in period t (i.e., regardless of whether a price review occurs in period t), because the firm has no costless memory.

The right-hand side of (1.4) defines a linear functional $T_{\Pi_t}[f_t]$ that maps any probability density f_t into another probability density f_{t+1} ; the subscript indicates that the mapping depends on the policy $\Pi_t \equiv (\Lambda_t, q_t^*)$. Given an initial prior f_0 and policies $\{\Pi_t\}$ for each of the periods $t \geq 0$, the law of motion (1.4) implies a sequence of priors $\{f_t\}$ for all periods $t \geq 1$. Note that if for any policy Π , the prior f is such that

$$T_{\Pi}[f] = f, \quad (1.5)$$

it follows that if a firm starts with the prior $f_0 = f$ and implements policy Π each period, the dynamics (1.4) imply that the firm will have prior $f_t = f$ in every period. Thus f is an *invariant distribution* for the Markov process describing the dynamics of q under this policy. In such a situation, we can say that the firm's prior each period corresponds to the long-run frequency with which different values of its normalized

price occur, under its constant policy Π . When the firm's prior is unchanging over time in this way, the constant prior makes it optimal for the firm to choose the same policy each period, which in turn makes it possible for the prior to remain constant. In the numerical analysis below, I shall be interested in computing statistics (for example, the frequency of price changes) for a stationary optimal plan of this kind.

The assumption that memory is (at least) as costly as information about current conditions external to the firm implies that under an optimal policy, the timing of price reviews is (stochastically) state-dependent, but *not* time-dependent, just as in full-information menu-cost models. When the cost θ of interim information is sufficiently large, the dependence of the optimal hazard on the current state is also attenuated, so that in the limit as θ becomes unboundedly large, the model approaches one with a constant hazard rate as assumed by Calvo (1983). If, instead, memory were costless, the optimal hazard under a stationary optimal plan would also depend on the number of periods since the last price review: there would be a different hazard function $\Lambda_n(q)$ for each value of n . In this case, in the limit of unboundedly large θ , each of the functions $\Lambda_n(q)$ would become a constant (there would cease to be dependence on q); but the constants would depend on n , and in the generic case, one would have Λ_n equal to zero for all n below some critical time, and Λ_n equal to 1 for all n above it. Thus the model would approach one in which prices would be reviewed at deterministic intervals, as in the analyses of Caballero (1989), Bonomo and Carvalho (2004), and Reis (2006). The analysis of this alternative case under the assumption of a finite positive cost of interim information is left for future work.

1.3 Stationary Optimal Price-Review Policies

We can now state the firm's dynamic optimization problem. Its dynamic price-review scheduling strategy is a sequence of policies $\{\Pi_t\}$ for each of the periods $t \geq 0$; given the initial prior f_0 , each such strategy implies a particular sequence of priors $\{f_t\}$ consistent with (1.4). The strategy is a *deterministic* sequence, insofar as in each period, the intended values of $\Lambda_t(q)$ and q_t^* depend only on t , and not on the signals received by the firm in any periods prior to t , on the timing of its price reviews prior to t , or on any information collected in the course of those reviews. This is because of the assumption that memory is costly; even if we imagine that the firm designs the signalling mechanism for period t and chooses its intended responses to signals in

period t only when that period is reached, it must solve this design problem — which allows it to *choose* how much memory to access in period t in making its price-review decision — without making use of any memory.¹⁶

The firm’s objective when choosing this strategy has three terms: the expected value of discounted profits (1.1), the expected discounted value of the costs of price reviews, and the discounted value of the costs of interim information used each period in that period’s price-review decision. The fixed cost of a price review is assumed to be $\kappa > 0$ in each period t in which such a review occurs; the cost of interim information is assumed to be θI_t in each period t (regardless of the signal received in that period), where I_t is the expected information used by a strategy that results in a hazard function $\Lambda_t(q)$, given the prior f_t for that period. In each case, the information costs are assumed to be in the same units as $\pi(q_t)$, and all costs in period t are discounted by the discount factor β^t .

A firm’s *ex ante* expected profits in any period t can be written as $\bar{\pi}(\Pi_t; f_t)$, where $\Pi_t = (\Lambda_t(q), q_t^*)$ is the policy followed in period t , f_t is the firm’s prior in period t (given its policies in periods prior to t), and

$$\bar{\pi}(\Pi; f) \equiv \int [\Lambda(q)\pi(q^*) + (1 - \Lambda(q))\pi(q)]f(q)dq.$$

The *ex ante* expected cost of price reviews in period t can be written as $\kappa\bar{\lambda}(\Pi_t; f_t)$, where

$$\bar{\lambda}(\Pi; f) \equiv \int \Lambda(q)f(q)dq$$

indicates the probability of a price review under a policy Π . Finally, the cost of interim information can be written (as in Woodford, 2008) as $\theta I_t = \theta\bar{I}(\Pi_t; f_t)$, where

$$\bar{I}(\Pi; f) \equiv \int \varphi(\Lambda(q))f(q)dq - \varphi(\bar{\lambda}(\Pi; f)), \quad (1.6)$$

¹⁶I assume here that a firm can implement a sequence of policies $\{\Pi_t\}$ which need not specify the same policy Π for each period t , without using “memory” of the kind that is costly. I assume that a firm has no difficulty remembering the strategy that it chose *ex ante*; what is costly is memory of *things that happen* during the execution of the strategy, that were not certain to happen *ex ante*. Note also that the firm’s price-review policy fails to be time-dependent, not because it lacks a “clock” to tell it the current value of t , but because it cannot costlessly remember whether it reviewed its pricing policy in any given previous period; it knows the value of t but not the value of n .

and $\varphi(\Lambda)$ is the Shannon binary entropy function, defined as

$$\varphi(\Lambda) \equiv \Lambda \log \Lambda + (1 - \Lambda) \log(1 - \Lambda) \quad (1.7)$$

in the case of any $0 < \Lambda < 1$, and at the boundaries

$$\varphi(0) = \varphi(1) = 0.$$

The firm's problem is then to choose a sequence of policies $\{\Pi_t\}$ for $t \geq 0$ to maximize

$$\sum_{t=0}^{\infty} \beta^t [\bar{\pi}(\Pi_t; f_t) - \kappa \bar{\lambda}(\Pi_t; f_t) - \theta \bar{I}(\Pi_t; f_t)], \quad (1.8)$$

where the prior evolves according to

$$f_{t+1} = T_{\Pi_t}[f_t] \quad (1.9)$$

for each $t \geq 0$, starting from a given initial prior f_0 . A *stationary optimal policy* is a pair (f, Π) such that if $f_0 = f$, the solution to the above dynamic problem is $\Pi_t = \Pi$ for all $t \geq 0$, and the implied dynamics of the prior are $f_t = f$ for all $t \geq 0$. Note that this definition implies that f satisfies the fixed-point relation (1.5), so that f is an invariant distribution under the stationary price-review policy Π .

1.4 A Recursive Formulation

The optimization problem stated above can be given a recursive formulation. This is useful for computational purposes, and also allows us to see how the problem involves a sequence of single-period price-review decisions of the kind analyzed in Woodford (2008). As a result, the characterization given there is both useful in computing the stationary optimal policy, and helpful in characterizing the random timing of price reviews of under such a policy.

For any initial prior f_0 , let $J(f_0)$ denote the maximum attainable value of the objective (1.8) in the problem stated above. Then standard arguments imply that $J(f)$ must satisfy a *Bellman equation* of the form

$$J(f_t) = \max_{\Pi_t} \{ \bar{\pi}(\Pi_t; f_t) - \kappa \bar{\lambda}(\Pi_t; f_t) - \theta \bar{I}(\Pi_t; f_t) + \beta J(f_{t+1}) \}, \quad (1.10)$$

where f_{t+1} is given by (1.9). If we can find a functional $J(f)$ (defined on the space of probability measures f) that is a fixed point of the mapping defined in (1.10),

then this is a *value function* for the optimization problem stated above. Moreover, the dynamic price-review scheduling problem can then be reduced to a sequence of single-period problems: in each period t , the policy Π_t is chosen to maximize the right-hand side of (1.10) subject to the constraint (1.9), given the prior f_t in the current period. The policy chosen each period then determines the prior in the next period through the law of motion (1.9). A stationary optimal policy is then a pair (f, Π) such that (i) if $f_t = f$, the solution to the problem (1.10) is $\Pi_t = \Pi$; and (ii) the distribution f is a fixed point (1.5) of the mapping defined by the policy Π .

This still does not make it easy to compute a stationary optimal policy, as one must first compute a functional $J(f)$ that is a fixed point of (1.10), and this is far from trivial, since (1.10) defines a mapping from a very high-dimensional function space into itself. Nor is the single-period policy problem defined in (1.10) as simple as the one considered in Woodford (2008). However, we can obtain an even simpler characterization by observing that $J(f_t)$ is necessarily a *concave* functional, that is furthermore differentiable at $f_t = f$ (the invariant distribution under the stationary optimal policy), so that for distributions f_t close enough to f , the value function can be approximated by a *linear* functional

$$J(f_t) \approx J(f) + \int j(q)[f_t(q) - f(q)]dq,$$

where $j(q)$ is an integrable function. (Note that the derivative function $j(q)$ is defined only up to an arbitrary constant, since $J(f_t)$ is not defined for perturbations of the set-valued function f_t that do not integrate to 1.) The concavity of $J(f_{t+1})$ then implies that $\Pi_t = \Pi$ solves the problem (1.10) when $f_t = f$ if and only if it solves the alternative problem

$$\max_{\Pi_t} \left\{ \bar{\pi}(\Pi_t; f) - \kappa \bar{\lambda}(\Pi_t; f) - \theta \bar{I}(\Pi_t; f) + \beta \int j(q)[f_{t+1}(q) - f(q)]dq \right\}, \quad (1.11)$$

where f_{t+1} is again given by (1.9).

Using (1.9) to substitute for f_{t+1} , the objective in (1.11) can alternatively be expressed as

$$(V(q_t^*) - \kappa) \int \Lambda_t(q)f(q) dq + \int V(q)(1 - \Lambda_t(q))f(q) dq - \theta \bar{I}(\Lambda_t; f), \quad (1.12)$$

where

$$V(q) \equiv \pi(q) + \beta \int j(\tilde{q})g(q - \tilde{q})d\tilde{q}, \quad (1.13)$$

and I have now written simply $\bar{I}(\Lambda_t; f)$, to indicate that the function \bar{I} defined in (1.6) does not depend on the choice of q^* . (Here the variable of integration q in (1.12) is the normalized price in period t *after* the period t disturbance to market conditions, but *before* the decision whether to conduct a price review. In (1.13), q is instead the normalized price that is charged, *after* any price review has occurred, while \tilde{q} is the normalized price in the following period, after that period's disturbance to market conditions, but before the decision whether to conduct a price review in that period.) Maximization of (1.12) is in turn equivalent to maximizing

$$\int L(q; q_t^*) \Lambda_t(q) dq - \theta \bar{I}(\Lambda_t; f), \quad (1.14)$$

if we define

$$L(q; q^*) \equiv V(q^*) - V(q) - \kappa. \quad (1.15)$$

Hence $\Pi_t = \Pi$ solves the problem (1.10) when $f_t = f$ if and only if it maximizes (1.14).

This, in turn, is easily seen to be true if and only if (i) q^* is the value of q that maximizes $V(q)$, and (ii) given the value of q^* , the hazard function Λ maximizes (1.14), which can alternatively be written as

$$\int [L(x) \Lambda(x) - \theta \varphi(\Lambda(x))] f(x) dx + \theta \varphi\left(\int \Lambda(x) f(x) dx\right). \quad (1.16)$$

As shown in Woodford (2008), the hazard function that maximizes (1.14) must satisfy the first-order condition

$$\equiv L(x) - \theta \varphi'(\Lambda(x)) + \theta \varphi'(\bar{\Lambda}) = 0 \quad (1.17)$$

almost surely, where

$$\bar{\Lambda} \equiv \int \Lambda(q) f(q) dq$$

is the average frequency of reviews of pricing policy. Thus each period a price-review policy Π_t is chosen that solves a single-period problem identical to the one considered in Woodford (2008), and in the case of a stationary optimal plan, this problem is the same each period. However, the definition of this problem involves the function $j(q)$; thus it may still seem necessary to solve the Bellman equation for the function $J(f)$.

In fact, though, we only need to know the derivative function $j(q)$. And an envelope-theorem calculation, differentiating (1.10) at $f_t = f$, yields

$$\begin{aligned}
j(q) &= \Lambda_t(q)\pi(q_t^*) + (1 - \Lambda_t(q))\pi(q) - \theta \left[\varphi(\Lambda_t(q)) - \varphi' \left(\int \Lambda_t(\tilde{q})f(\tilde{q})d\tilde{q} \right) \right] \\
&\quad - \kappa\Lambda_t(q) + \beta \int j(\tilde{q})[\Lambda_t(q)g(q_t^* - \tilde{q}) + (1 - \Lambda_t(q))g(q - \tilde{q})]d\tilde{q} \\
&= \Lambda_t(q)[V(q_t^*) - \kappa] + (1 - \Lambda_t(q))V(q) - \theta \left[\varphi(\Lambda_t(q)) - \varphi' \left(\int \Lambda_t(\tilde{q})f(\tilde{q})d\tilde{q} \right) \right] \\
&= V(q) + \Lambda_t(q)L(q; q_t^*) - \theta \left[\varphi(\Lambda_t(q)) - \varphi' \left(\int \Lambda_t(\tilde{q})f(\tilde{q})d\tilde{q} \right) \right] \\
&= V(q) - \theta[\varphi(\Lambda_t(q)) - \varphi'(\Lambda_t(q))\Lambda_t(q)] \\
&= V(q) - \theta \log(1 - \Lambda_t(q)).
\end{aligned}$$

Here the second line uses the definition (1.13) of $V(q)$; the third line uses the definition (1.15) of $L(q; q^*)$; the fourth line uses the fact noted above that a solution to the problem (1.11) — and accordingly, a solution to the problem (1.10) — must satisfy the first-order condition (1.17) to substitute for $L(q; q^*)$; and the final line uses the definition (1.7) of the binary entropy function $\varphi(\Lambda)$. Note also that on each line, I have suppressed an arbitrary constant term, since $j(q)$ is defined only up to a constant.

Substituting the above expression for $j(q)$ into the right-hand side of (1.13), we obtain

$$V(q) \equiv \pi(q) + \beta \int [V(\tilde{q}) - \theta \log(1 - \Lambda(\tilde{q}))] g(q - \tilde{q})d\tilde{q}, \quad (1.18)$$

a fixed-point equation for the function $V(q)$ that makes no further reference to either the value function J or its derivative. A stationary optimal policy then corresponds to a triple (f, Π, V) such that (i) given the policy Π , the function V is a fixed point of the relation (1.18); given the pseudo-value function V and the prior f , the policy Π is such that q^* maximizes V and Λ maximizes (1.16); and (iii) given the policy Π , the distribution f is an invariant distribution, *i.e.*, a fixed point of relation (1.5).

This characterization of a stationary optimal policy reduces our problem to a much more mathematically tractable one than solution of (1.10) for the value function $J(f)$. We need only solve for two real-valued functions of a single real variable, the functions $V(q)$ and $\Lambda(q)$; a probability distribution $f(q)$ over values of that same single real variable; and a real number q^* . These can be solved for using standard methods of function approximation and simulation of invariant distributions, of the kind discussed for example in Miranda and Fackler (2002).

2 A Model of Monopolistically Competitive Price Adjustment

Let us now numerically explore the consequences of the model of price adjustment developed in section 1, in the context of an explicit model of the losses from infrequent price adjustment of a kind that is commonly assumed, both in the literature on canonical (full-information) SDP models and in ET models of inflation dynamics. The economy consists of a continuum of monopolistically competitive producers of differentiated goods, indexed by i , and in the case considered in this section, I shall assume that the only shocks to the economy are good-specific shocks, so that there is no aggregate uncertainty.

Let us suppose that each household seeks to maximize a discounted objective

$$E_0 \sum_{t=0}^{\infty} \beta^t \left[\frac{C_t^{1-\sigma^{-1}}}{1-\sigma^{-1}} - \lambda \frac{H_t^{1+\nu}}{1+\nu} \right]$$

where

$$C_t \equiv \left[\int_0^1 (c_t(i)/A_t(i))^{\frac{\epsilon-1}{\epsilon}} di \right]^{\frac{\epsilon}{\epsilon-1}} \quad (2.1)$$

is an index of consumption of the various differentiated goods indexed by i , $A_t(i)$ is a good-specific shock to preferences, H_t is the hours of labor supplied by the household (in the sector-specific labor market in which the particular household works), and the preference parameters satisfy $0 < \beta < 1$, $\sigma, \lambda > 0$, $\nu \geq 0$, and $\epsilon > 1$. Each differentiated good is supplied by a monopolistically competitive firm, with production function

$$y_t(i) = A_t(i)h_t(i)^{1/\phi}.$$

Here $h_t(i)$ indicates the hours of labor employed (of the specific type required for the production of good i), $A_t(i)$ is a good-specific productivity factor, and $\phi \geq 1$. Note that the good-specific factor $A_t(i)$ is assumed to shift *both* the relative preference for and the relative cost of producing good i . The assumption that the idiosyncratic shock shifts both preferences and technology in this way is plainly a very special (and rather artificial) one, but it has the advantage of making the profits of a firm i a function only of a single variable, the normalized price of good i (defined below);¹⁷ this is convenient not only because it makes the firm's problem a "tracking problem" of exactly the kind

¹⁷Another convenient feature of this assumption is that each firm's profit function is shifted in

assumed by Caballero and Engel (1993a, 1993b), but because computation is simpler in the case of a model with a one-dimensional state space. (The analysis proposed here can easily be extended to the case of separate good-specific shocks to preferences and technology, with only a modest increase in numerical complexity.)

Except for the assumption of the good-specific shocks $\{A_t(i)\}$, the model is exactly like the New Keynesian model of monopolistically competitive price adjustment expounded in Woodford (2003, chap. 3). Each household is assumed to own an equal share of each of the firms, and households' idiosyncratic income risk (owing to having specialized in supplying labor to a particular sector) is perfectly shared through insurance contracts, so that households' budgets are all identical in equilibrium, and their state-contingent consumption plans as well. As a consequence, each household values random income streams in the same way, and the (nominal) stochastic discount factor is given by

$$Q_{t,T} = \beta^{T-t} \left(\frac{C_t}{C_T} \right)^{\sigma^{-1}} \frac{P_t}{P_T},$$

where C_t is aggregate (and each household's individual) consumption of the composite good (defined in 2.1) in period t , and P_t is the price of a unit of the composite good. Each firm i sets prices so as to maximize the discounted value of its profits using this stochastic discount factor. This is equivalent to maximization of an objective of the form

$$E_0 \sum_{t=0}^{\infty} \beta^t \tilde{\Pi}_t(i), \quad (2.2)$$

where $\tilde{\Pi}_t(i) \equiv \Pi_t(i) C_t^{-\sigma^{-1}} / P_t$, and $\Pi_t(i)$ is the nominal profit (revenue in excess of labor costs) of firm i in period t .

The Dixit-Stiglitz preference specification (2.1) implies that each firm i faces a demand curve of the form

$$y_t(i) = A_t(i)^{1-\epsilon} Y_t \left(\frac{P_t(i)}{P_t} \right)^{-\epsilon} \quad (2.3)$$

for its good, where $P_t(i)$ is the price of good i , Y_t is production of (and aggregate the same way by a shock to aggregate nominal expenditure (considered in section 4 below) as it is by the idiosyncratic shock; again the profit function is unchanged once the price is appropriately renormalized.

demand for) the composite good defined in (2.1), and the price index

$$P_t \equiv \left[\int (P_t(i) A_t(i))^{1-\epsilon} di \right]^{\frac{1}{1-\epsilon}} \quad (2.4)$$

indicates the price of a unit of the composite good. Optimal labor supply in each sectoral labor market implies that the real wage paid by firm i will equal $W_t(i)/P_t = \lambda h_t(i)^\nu C_t^{\sigma-1}$. Using this expression for the wage and using the production function to solve for the labor demand of firm i as a function of its sales, one finds that the profits of firm i in period t are equal to

$$\begin{aligned} \Pi_t(i) &= P_t(i) y_t(i) - W_t(i) h_t(i) \\ &= P_t(i) y_t(i) - \lambda P_t C_t^{\sigma-1} \left(\frac{y_t(i)}{A_t(i)} \right)^\eta, \end{aligned}$$

where $\eta \equiv \phi(1 + \nu) \geq 1$.

If we define the *normalized price* of firm i as

$$Q_t(i) \equiv \frac{P_t(i) A_t(i) \bar{Y}}{P_t Y_t},$$

where $\bar{Y} > 0$ is a normalization factor (chosen below), and let \mathcal{Q}_t be the population distribution of values of $Q_t(i)$ across all firms in period t , then it follows from (2.4) that

$$Y_t = Y(\mathcal{Q}_t) \equiv \bar{Y} \left[\int Q_t(i)^{1-\epsilon} di \right]^{-1/(1-\epsilon)}.$$

One can then write the period contribution to the objective (2.2) of firm i as a function solely of the firm's normalized price $Q_t(i)$ and the distribution of normalized prices \mathcal{Q}_t ,

$$\tilde{\Pi}_t(i) = \bar{Y}^{\epsilon-1} Y(\mathcal{Q}_t)^{2-\sigma-1-\epsilon} Q_t(i)^{1-\epsilon} - \lambda \bar{Y}^{\eta\epsilon} Y(\mathcal{Q}_t)^{\eta(1-\epsilon)} Q_t(i)^{-\eta\epsilon},$$

using the demand function (2.3) to solve for the firm's sales, and the fact that $C_t = Y_t = Y(\mathcal{Q}_t)$ in equilibrium.

In the equilibrium discussed in the next section, I assume that there are only idiosyncratic shocks, and consider only a *stationary equilibrium* in which the population distribution of normalized prices is constant over time. In such an equilibrium, the decision problem of each individual firm involves an objective of the form (1.1), where I now define $q_t(i) \equiv \log Q_t(i)$. (Note that this is consistent with the definition in section 1, if we define $m_t(i) \equiv \log(P_t Y_t / A_t(i) \bar{Y})$). Note also that, as assumed in

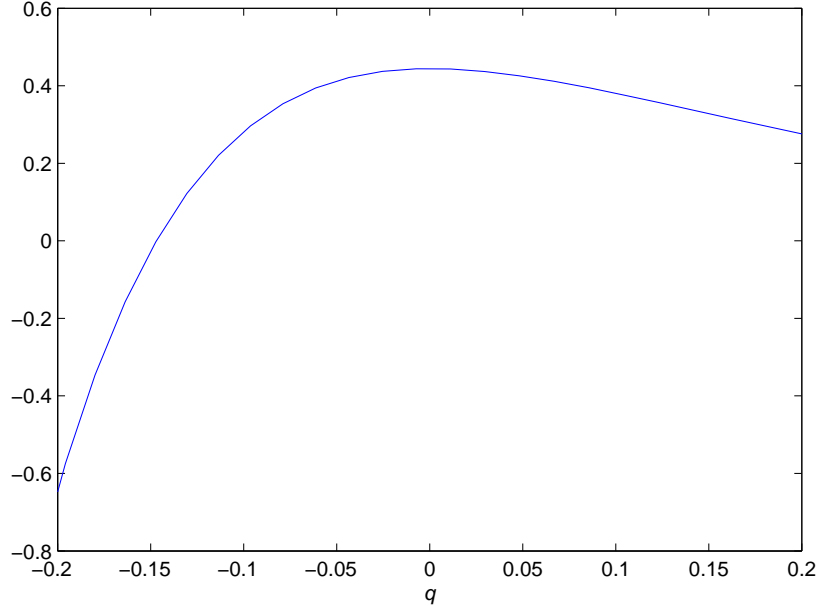


Figure 1: The normalized profit function $\pi(q)$.

section 1, the evolution of $\{m_t(i)\}$ is outside the control of firm i .) It is convenient, in reporting the numerical results below, to choose the normalization factor¹⁸ \bar{Y} so that

$$\bar{Y} = \left[\frac{Y(\bar{Q})^{(\eta-1)(1-\epsilon)+\sigma^{-1}-1}}{\lambda\mu\eta} \right]^{\frac{1}{1+(\eta-1)\epsilon}},$$

where \bar{Q} is the stationary distribution of normalized prices and $\mu \equiv \epsilon/(\epsilon - 1) > 1$ is the desired markup factor (under full information) of a firm facing a demand curve of the form (2.3). In this case, the firm's objective can be written as (1.1), where the period profit function is given by¹⁹

$$\pi(q) \equiv Q^{1-\epsilon} - \frac{1}{\eta\mu} Q^{-\eta\epsilon}. \quad (2.5)$$

The profit function is thus a member of a two-parameter family of functions

¹⁸This is essentially a choice of the units in which output is measured. The point of the proposed normalization is to make the profit-maximizing normalized price equal to 1.

¹⁹This expression for the profit function omits a positive multiplicative factor that is time-invariant under the assumptions stated in the text. The normalization has no consequences for the optimal price-review scheduling problem.

(depending only on the values of ϵ and η). Under our maintained assumptions that $\epsilon > 1$ and $\eta \geq 1$, this is necessarily a single-peaked function, reaching its maximum at the value $q = 0$. Moreover, the profit function is asymmetric, with profits falling more sharply for negative values of q than for positive values of the same size, as shown in Figure 1. This asymmetry results in an asymmetric hazard function for the price-review decision, as illustrated in the next section.

For simplicity (*i.e.*, to make the state space as small as possible), I also assume that the firm-specific factor follows a random walk,

$$\log A_t(i) = \log A_{t-1}(i) + z_t(i),$$

where the innovation $z_t(i)$ is drawn independently (both across firms i and across time periods t) each period from a distribution $g(z)$. In section 3 I also assume that aggregate nominal expenditure $P_t Y_t$ is constant over time (though the consequences of a disturbance to aggregate expenditure are discussed in section 4). Under this assumption, the normalized price can be written as $q_t(i) = p_t(i) - m_t(i)$, where $m_t(i)$ evolves as in (1.2).

Under these assumptions, each firm's price-review scheduling problem is a problem of the form discussed in section 1. We can then characterize the stationary optimal price-review policy for each firm as in that section. Associated with the stationary optimal policy is a long-run frequency distribution $f(q)$ for each firm's normalized price $q_t(i)$, and given the independence across firms of the evolution of $q_t(i)$,²⁰ this distribution $f(q)$ is also the (time-invariant) population distribution \bar{Q} of normalized prices across firms. Note that under the normalization of $q_t(i)$ proposed, it is not necessary to solve for the population distribution in order to solve the individual firm's price-review scheduling problem; however, we do need to solve for \bar{Q} in order to determine the (time-invariant) level of equilibrium output in the stationary equilibrium, and the value of the normalization factor \bar{Y} , in order to interpret the dynamics of normalized prices in terms of their implications for non-normalized prices and quantities.²¹

²⁰Both the innovations $z_t(i)$ in the firm-specific factor and the randomness of the signalling mechanism used to schedule price reviews (reflected in the hazard function $\Lambda(q)$) are assumed to be independent across firms i .

²¹Solution for the optimal price-review policy does require that one specify the information-cost parameters κ and θ in the same units as the normalized profit function $\pi(q)$, and so if one knows

3 Dynamics of Individual Prices under the Stationary Optimal Policy

I now illustrate the implications of the model set out in section 2 for the dynamics of individual price adjustment to firm-specific disturbances, by numerically computing the stationary optimal policy under calibrated parameter values that allow the model to match (at least roughly) certain statistics of microeconomic data sets. In addition to the functional-form assumptions already stated in section 2, I assume that the distribution $g(z)$ is $N(0, \sigma_z^2)$. I compute the stationary optimal policy using an algorithm of the kind described in the Appendix, using numerical parameter values $\beta = 0.9975$ (corresponding to a 3 percent annual rate of time preference, on the understanding that model “periods” represent months), $\epsilon = 6$ (implying a desired markup of 20 percent), and $\eta = 1.5$ (consistent with the values $\phi = 1.5$, corresponding to an elasticity of output with respect to the labor input of $2/3$, and $\nu = 0$, as in a model with “indivisible labor”).

The remaining three parameters that must be assigned numerical values — σ_z , κ , and θ — are chosen so that the model’s numerical predictions are at least roughly consistent with microeconomic evidence on individual price changes, as discussed below in section 3.2. The predictions for three empirical statistics (shown in Table 3) — the average frequency of price changes, the mean size of price change (mean absolute value of the change in log price), and the coefficient of kurtosis of the distribution of price changes — are used to select realistic values for the three parameters. The baseline parameter values selected on this ground are $\sigma_z = 0.06$ (a one-standard-deviation firm-specific shock changes the firm’s optimal price by 6 percent), $\kappa = 0.5$ (the cost of a price review is half a month’s output), and $\theta = 5$.

In addition to the results for this value of θ , I also present numerical results for a variety of other values of θ , listed in Table 1, including values both larger and smaller

the value of these costs in real (non-normalized) terms, one cannot determine the values of κ and θ required for the calculations described in section 1 without having determined \bar{Q} . But if, as below, one infers the values of κ and θ from their implications for the size and frequency of price changes, rather than from any direct evidence on the size of these costs, then it is not necessary to have already determined \bar{Q} in order to compute the implications of particular values of κ and θ for those statistics. One does have to solve for the implied distribution \bar{Q} in order to determine what these values of κ and θ correspond to in terms of non-normalized costs.

Table 1: Resource expenditure on information, for values of θ . Each share is measured in percentage points.

θ	s_κ	s_θ	r_θ
0	5.0	0	100
.05	4.5	1.16	56.7
.5	5.2	2.72	9.7
5	9.2	2.11	0.78
50	12.5	0.70	0.03
∞	15.0	—	0

than the baseline value.²² Because the assumption of a finite positive cost θ of interim information is the main novelty of the model presented here, it is of particular interest to explore the consequences of alternative values of this parameter.

Table 1 provides an indication of the magnitude of information costs implied by various values of θ , showing in each case the implied cost to the firm of inter-review information collection (*i.e.*, the cost of the information on the basis of which decisions are made about the scheduling of price reviews), as well as the cost to the firm of price reviews themselves, both as average shares of the firm’s revenue. (These two shares are denoted s_θ and s_κ respectively.) The table also indicates how the assumed information used by the firm in deciding when to review its prices compares to the amount of information that would be required in order to schedule price reviews optimally; the information used is fraction r_θ of the information that would be required for a fully optimal decision, *given* the firm’s value function for its continuation problem in each period (which depends on the fact that, at least in the future, it does *not* expect to schedule price reviews on the basis of full information). A value of $\theta = 5$, for example, might seem high, in that it means that the cost per nat of information²³ is 5 months of the firm’s steady-state revenue. However, under the stationary optimal

²²The bottom line of the table describes limiting properties of the stationary optimal plan, as the value of θ is made unboundedly large, *i.e.*, in the “Calvo limit”.

²³A “nat” is equal to 1.44 bits (binary digits) of information. The quantity of information is measured in nats in this paper, as I use natural logarithms (rather than base 2 logarithms) in defining the entropy measure.

policy, the firm only uses information each month in deciding whether to review its pricing policy with a cost equivalent to about two percent of steady-state revenue. And since this is 0.78 percent of the information that would be required to make a fully optimal decision (as assumed in standard SDP models), this specification of the information cost implies that it would cost less than three times total revenues for the firm to make a fully optimal decision each month.²⁴ This is a substantial cost, but perhaps not an unrealistic one; firms surely would find it prohibitively expensive to be constantly well-enough informed to make a precisely optimal decision each month about the desirability of a price review.²⁵

An information cost of $\theta = 5$ seems high when expressed as a cost per nat (or cost per bit) of information, because I allow the signal s to be designed so as to focus on precisely the information needed for the manager’s decision; once I have done so, one can only explain imprecision in the decisions that are taken under the hypothesis that the information content of s must be quite small, or alternatively, that the marginal cost of increasing the information content of the signal s is quite high.²⁶

The value assumed for κ (half a month’s output) may also seem large. However, under the baseline calibration (the case $\theta = 5$ in Table 1), this only implies that the costs of reviews of pricing policy are about 9 percent of the value of the firm’s output. In the firm studied by Zbaracki *et al.* (2004), the total managerial costs of reviews of pricing policy are reported to be only 1.4 percent of total operating costs. However, the total costs of price changes (counting also physical “menu costs” and the costs of communication of the new prices to customers) are reported to account for more than 6 percent of operating costs, and in the present model all of these costs

²⁴Here I refer to the cost of making a fully optimal decision in one month only, taking for granted that one’s problem in subsequent months will be the information-constrained problem characterized here, and not to the cost of making a fully optimal decision each month, forever. In Table 1, the information cost of a fully optimal decision is computed using the value function $V(q)$ associated with the stationary optimal policy corresponding to the given value of θ .

²⁵Under a more realistic calibration of the model, “periods” should perhaps correspond to weeks rather than to months, and this would doubtless affect the calculations reported in Table 1. Here I interpret the model “periods” as months because the micro data discussed in section 3.2 are monthly.

²⁶It is important to understand that the parameter θ does not represent a cost-per-letter of having a staff member read the *Wall Street Journal*; it is instead intended to represent a cost of getting the attention of the manager who must make the decision, once the staff have digested whatever large amount of information may have been involved in the preparation of the signal s that must be passed on to the manager.

Table 2: The optimal value of q^* for alternative values of θ .

θ	q^*
0	.015
.05	.015
.5	.039
5	.068
50	.067
∞	.076

are represented by the parameter κ . Thus under the baseline calibration used here, the total costs involved in reviews of pricing policy are only one and one-half times as large as those at the firm studied by Zbaracki *et al.*²⁷

3.1 The Stationary Optimal Policy under Alternative Costs of Information

The optimal price-review policy of an individual firm is completely specified by the reset value for the normalized price, q^* , and the hazard function $\Lambda(q)$. (An advantage of the univariate case considered here is that the hazard is a function of a single real variable, and can easily be plotted.) Table 2 shows the optimal value of q^* for a range of values for the information cost θ , and Figure 2 plots the corresponding optimal hazard functions. One observes that q^* is positive, and by several percentage points in the case of the larger values of θ . The optimal reset value is positive, even though it would be optimal to set $q_t(i) = 0$ at all times in the absence of information costs, due to the asymmetry of the profit function seen in Figure 1. Because the losses associated with a price that is too low are greater than those associated with a price that is too high by the same number of percentage points, it is prudent to set one's price slightly higher than one would if one expected to be able to adjust the price again immediately in the event of any change in market conditions, in order to reduce the probability of having a price that is too low. The size of the bias that is optimal is greater when interim information is costly, but it is still positive even when $\theta = 0$;

²⁷The reason for choosing a value of κ of this size is explained below in section 3.2.

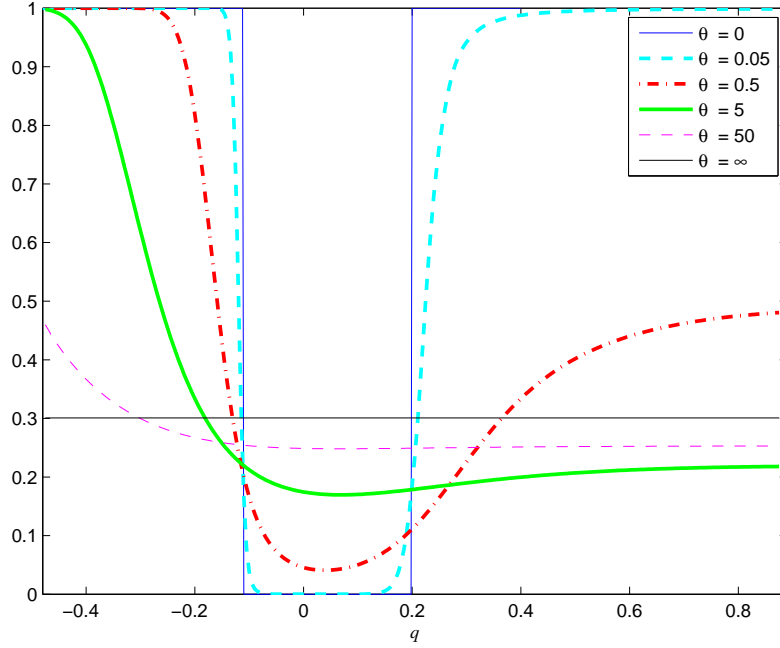


Figure 2: The optimal hazard function $\Lambda(q)$, for alternative values of θ .

for even in that case, prices are not re-optimized continually, owing to the cost κ of price reviews.

In the case that $\theta = 0$, the optimal hazard function has the “square well” shape associated with standard SDP models: there is probability 0 of adjusting inside the S s thresholds, and probability 1 of adjusting outside them. For positive values of θ , one instead has a continuous function taking values between 0 and 1, with the lowest values in the case of “price gaps” near zero,²⁸ and higher values in the case of large price gaps of either sign. When θ is small (though positive), as in the case $\theta = 0.05$ shown in the figure, the hazard function is still barely above 0 for small price gaps, and rises rapidly to values near 1 for price gaps that are only a small distance outside the “zone of inaction” under full information. But for larger values of θ , the optimal hazard function is significantly positive even for price gaps near zero, and increases only slightly even for price gaps far outside the full-information “zone of inaction”.

²⁸Here I define the “price gap” as the value of $q_t(i) - q^*$, i.e., the gap between the current log price and the price that would currently be chosen in the event of a review.

Note that the prediction of a substantial positive hazard rate even when the price gap is zero is consistent with the empirical evidence of Eichenbaum *et al.* (2008), who find a positive probability of adjustment of a good’s “reference price” (greater than 15 percent) even in a quarter in which the normalized reference price (identified in their case by the markup of the reference price over the cost of the good to the retailer) is equal to the long-run average normalized reference price.²⁹

It is also interesting to observe that for substantially positive (but still finite) values of θ , the optimal hazard function is quite asymmetric: the probability of a price review rises much more rapidly in the case of negative price gaps than in the case of positive price gaps of the same magnitude. In fact, this asymmetry is of the same sign as has been found to best fit U.S. data on both aggregate inflation dynamics (Caballero and Engel, 1993a) and on the distribution of individual price changes (Caballero and Engel, 2006). Eichenbaum *et al.* (2008) also provide evidence for asymmetry of the same sign in the hazard function for adjustment of the “reference price” of a good as a function of its normalized price (markup over cost). The present model provides a theoretical explanation for the asymmetry for which these authors argue on empirical grounds.

In the limit as θ is made unboundedly large, the optimal hazard approaches a positive value between zero and one that is independent of the size of the price gap. (This limiting hazard is shown by the horizontal solid line in the figure, at a hazard of approximately 0.3.) The convergence can be seen particularly clearly in the case of price gaps near zero, where the hazard approaches the limiting value monotonically from below; convergence similarly occurs for price gaps outside the full-information S s thresholds, though in this case the convergence is non-monotonic. Thus in the limit of very costly interim information, the model predicts behavior like that of the Calvo model (in which the hazard is a constant positive rate regardless of the price gap). In fact, already in the case $\theta = 50$, one can see from the figure that the hazard rate is essentially constant for a large range of normalized prices, corresponding to price gaps anywhere between negative 20 percent³⁰ and positive 80 percent or more.

²⁹See their Figure 12, panel A. Eichenbaum *et al.* also find (Figure 12, panel C) that when a good’s reference price is adjusted, the normalized price (or markup) is reset to the same value — corresponding to the long-run average normalized price — regardless of the value of the normalized price prior to the adjustment, as predicted by the model here.

³⁰Note that since (according to Table 2) q^* is equal to 7.6 percent in this case, a price gap of -20 percent corresponds to the value $q = -.124$ in the figure.

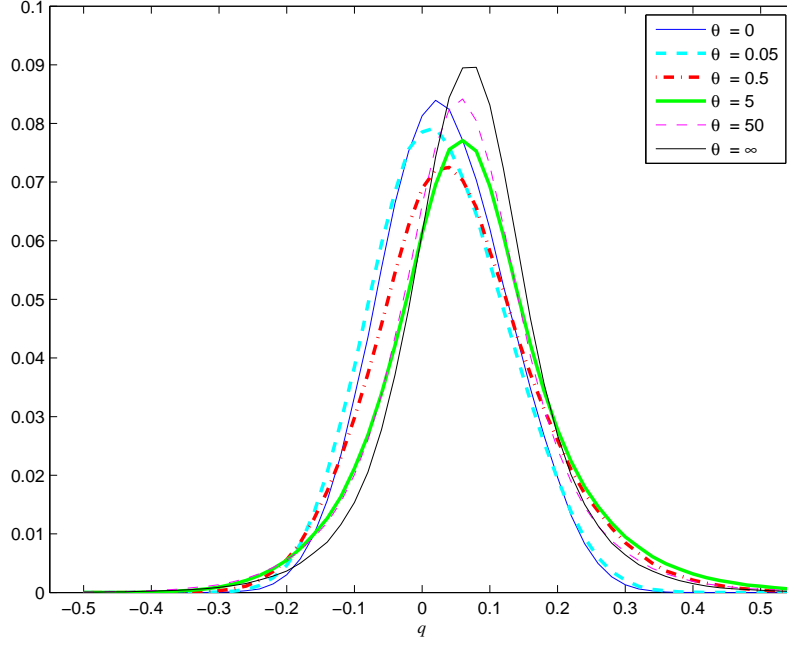


Figure 3: The invariant distribution $f(q)$, for alternative values of θ .

The invariant distribution $f(q)$ implied by the optimal policy $(\Lambda(q), q^*)$ is shown in Figure 3 for each of these same values of θ . The tightness of the range of variation in normalized prices is fairly similar across the different values of θ ; the main difference is that the distributions are shifted slightly to the right in the case of larger values of θ , because of the larger positive values of q^* in these cases. In the case of low values of θ (for example, as one moves from $\theta = 0.05$ to $\theta = 0.5$), increasing θ increases slightly the range of variation in the normalized price, as the hazard no longer rises toward one quite so sharply in the case of price gaps outside the full-information Ss thresholds; but as θ increases further, the range of variation shrinks again, owing to the increase in the hazard rate in the case of price gaps within the full-information Ss thresholds.

Perhaps the most important implication of Figure 3 is that regardless of the size of θ , a firm's normalized price will much of the time be within an interval between (roughly) -0.15 and +0.30. This means that it is only the hazard function over that interval that matters very much for the equilibrium dynamics of prices; in particular,

the degree to which an appropriately calibrated Calvo model will approximate the equilibrium dynamics depends only on the constancy of the hazard function over that interval. From this we can see that the Calvo model will be quite a good approximation in the case of a value of θ equal to 50 or higher. But even for a value on the order of $\theta = 5$ (our baseline value), the hazard function is fairly constant over most of this interval, and so the Calvo model should not be too bad an approximation in this case either.

3.2 Comparison with Data on Individual Price Changes

We have seen that, depending on the value of θ , the present model predicts dynamics of price adjustment like those of an Ss model, like those of the Calvo model, or like those of a “generalized Ss” model of the kind discussed by Caballero and Engel (1993a, 1993b). This raises an obvious question: which of these alternative parameterizations of the model is more empirically realistic? One way of answering the question is to compare the extent to which the model’s predictions under alternative parameterizations match the properties of individual price changes in microeconomic data sets, as summarized in studies like those of Bils and Klenow (2004), Klenow and Kryvtsov (2008), Midrigan (2008), or Nakamura and Steinsson (2008a).

Table 3 displays the values for several key statistics that are predicted by the calibrated model discussed above, in the case of several alternative values of θ . The first column indicates the value of $\bar{\Lambda}$, the average frequency of price reviews (and hence of price changes). The second column indicates the mean value of the absolute price change $|\Delta p_t(i)|$ in those months in which a price review (and hence a price change) occurs. The third column indicates the fraction of price changes that are smaller than 0.05 in absolute value (again, conditional on a price change occurring), and the fourth column reports the coefficient of kurtosis for the distribution of price changes (conditional on a price change occurring). One observes that, for given values of the other model parameters, increasing the value of θ increases the average frequency of price changes (for values of θ greater than 0.5), reduces the average size of price changes, increases the frequency of small price changes, and increases the kurtosis of the distribution of price changes.

The final two rows of the table indicate two different sets of “target” values for our calibration, deriving from two different types of microeconomic data sets. The

Table 3: Predictions regarding the size and frequency of price changes, for alternative values of θ . (*Probability that $|\Delta p| < .06$.)

θ	$\bar{\Lambda}$	$E[\Delta p]$	$Pr(\Delta p < .05)$	$K(\Delta p)$
<i>Predictions of calibrated model</i>				
0	.100	.186	0	1.2
.05	.090	.195	.000	1.3
.5	.104	.160	.145	2.2
5	.184	.105	.341	4.1
50	.250	.089	.397	4.8
∞	.301	.081	.421	4.9
<i>BLS data: Klenow-Kryvtsov (2008) ‘like prices’</i>				
	.168	.118	.427	
<i>Midrigan (2008) calibration ‘targets’</i>				
	.24	.12	.28*	4.0

first set of targets are based on the BLS data on individual prices underlying the CPI. The statistics reported are those given by Klenow and Kryvtsov (2008) for what they call “like price” changes, in which a “regular price” for a good is compared to the previous “regular price” for the good and a “sale price” is compared to the previous “sale price” for that good, but sale prices and regular prices for the same good are treated as if they are from price series for separate goods.³¹ (Klenow and Kryvtsov do not report a value for the kurtosis.) The second set of targets are the calibration targets proposed by Midrigan (2008), on the basis of a summary of evidence from two different sets of scanner data.

It is evident from the table that, given the assumed values of the other parameters,

³¹I use these statistics as the target for calibration of the model, under the interpretation that those goods for which frequent “sale prices” are observed have a pricing policy that involves both a regular price and a sale price; a review of the pricing policy is indicated when the regular price or the sale price is changed, but not when the price of the good switches from the regular price to the sale price. If one were instead to count as price changes only changes in “regular prices,” the statistics reported by Klenow and Kryvtsov would be $\bar{\Lambda} = 0.139$, $E[|\Delta p|] = 0.113$, and $Pr(|\Delta p| < .05) = 0.443$. Thus the only material difference would be the lower frequency of price changes if only changes in regular prices are counted.

a value of θ on the order of 5 is most consistent with the microeconomic evidence. A value of θ substantially smaller than this would imply less frequent price changes, a larger average size of price changes, and too little kurtosis relative to the micro data sets; but a value of θ larger than this would imply more frequent price changes, of smaller average size, than in the micro data sets, as well as more small price changes (and a somewhat higher coefficient of kurtosis) than Midrigan (2008) reports.

The statistics that are most directly revealing regarding the empirically realistic value of θ are the coefficient of kurtosis and the frequency of small price changes. When $\theta = 0$, there are no small price changes at all, because every price change is a movement from a normalized price outside the Ss thresholds to the optimal normalized price q^* .³² For small positive values of θ , small price changes continue to be extremely infrequent. But the fraction of small price changes increases steadily as θ is increased. At the baseline value of the information cost parameter ($\theta = 5$), the model predicts that price changes smaller than 5 percent should occur with a frequency that is still not as large as the frequency reported by Klenow and Kryvtsov for the BLS data, but that is larger than the frequency reported by Midrigan (2008) for his scanner data sets.³³ Midrigan’s statistic is arguably more relevant for our purposes, since he corrects for heterogeneity in the size of price changes in the case of different types of goods; it is also more consistent with the findings of Eichenbaum *et al.* (2008), who report that 26 percent of “reference price” changes are of 5 percent or smaller (see their Figure 15). But even a target of 0.26 or 0.28 would suggest that a value of θ substantially larger than 0.5 is needed.

Nor is the increased fraction of price changes that are smaller than 5 percent for larger values of θ a consequence solely of the smaller average size of price changes; there is also an increased frequency of prices that are small *relative* to the mean absolute price change. Figure 4 displays the complete (stationary) distribution of

³²This does not mean that it is not possible to have price changes of less than 5 percent; by choosing a value of κ that is substantially smaller, it is possible to calibrate the model so that the Ss thresholds are much closer to q^* than for the parameters used here. But this would allow us to explain the existence of a substantial fraction of price changes smaller than 5 percent only by making the mean absolute price change much smaller than what is observed in the micro data sets.

³³The frequency reported on the bottom line of Table 3, 28 percent, is actually the value that Midrigan reports for the probability of a price change of less than half the mean size, meaning less than .06. The frequency of price changes less than .05 in size would be somewhat smaller than this (but is not reported by Midrigan).

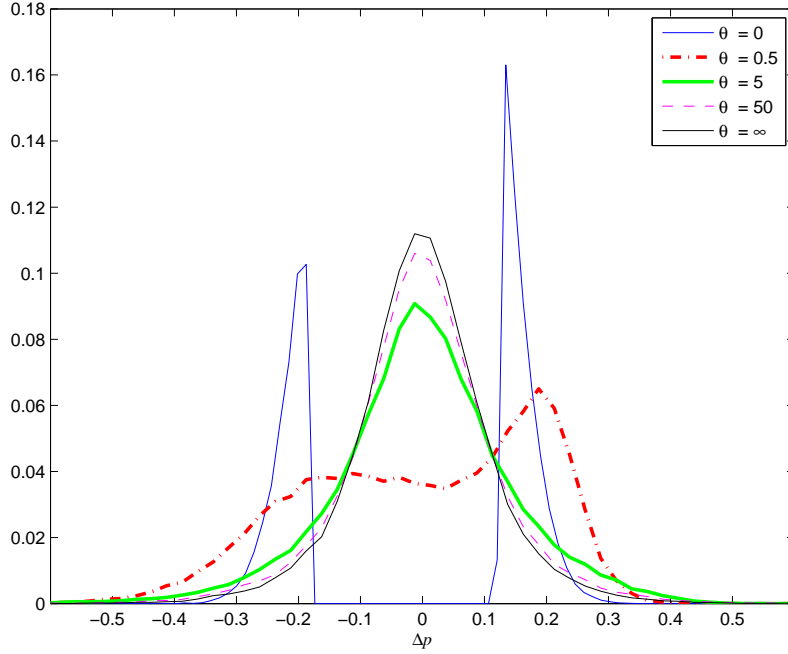


Figure 4: The (unconditional) distribution of individual price changes, for alternative values of θ .

price changes for each value of θ ; the increased probability mass for small values of q as θ increases is clearly apparent. The distribution of price changes predicted by the model when θ is small is quite different from the empirical distributions shown in Midrigan (2008, Figure 1). The distributions found in the microeconomic data sets analyzed by Midrigan are unimodal and leptokurtic, with a higher peak and fatter tails than a normal distribution with the same variance. In the present model, for all low enough values of θ the distribution is bi-modal (as in a standard Ss model); for somewhat higher values of θ , it is unimodal but platykurtic (with a flat peak and thin tails); and it becomes leptokurtic only for high enough values of θ . In the case of the baseline values of the other model parameters, the distribution is leptokurtic only if θ is on the order of 5 (as assumed in the baseline calibration) or higher.

On the basis of the microeconomic evidence, Midrigan suggests a target value of 4 for the coefficient of kurtosis.³⁴ The last column of Table 3 shows that the present

³⁴Klenow and Kryvtsov do not report a value for this statistic in the BLS data. However, their

model predicts a coefficient of kurtosis ($K[|\Delta p|]$) of roughly this magnitude when θ is equal to 5,³⁵ again suggesting that this is the most empirically realistic magnitude to assume for the interim information costs. Changing the value of other parameters (such as κ) can change the size distribution of price changes, but does not change the prediction that the distribution is leptokurtotic only for large enough values of θ ; and while the θ that is “large enough” could be much smaller if we were to assume a much smaller value of κ , it would remain the case that a value of θ large enough to make the distribution leptokurtic is large enough to imply price adjustment dynamics substantially different from those of a standard menu-cost (Ss) model (the $\theta = 0$ case), and instead fairly similar to those implied by a Calvo model (the $\theta \rightarrow \infty$ limit).

3.3 Calibration of σ_z and κ

Of course, inference about the realistic value of θ from Table 3 depends on accepting that the values assigned to the other model parameters in Table 3 are realistic. The values of σ_z and κ in particular deserve further discussion, as these are chosen not on the basis of any direct measures of these two quantities, but rather on the basis of the resulting predictions for the statistics reported in Table 3.

In fact, we can determine appropriate values for each of the three parameters σ_z, κ and θ on the basis of the target values for the statistics listed in Table 3, as long as variation in each of the three parameters has sufficiently distinct effects on the predicted statistics. It suffices to pick targets for three statistics that are affected by the parameters in sufficiently different ways, namely, for the average frequency of price changes $\bar{\Lambda}$, the average size of price changes $E[|\Delta p|]$, and the coefficient of kurtosis $K[|\Delta p|]$. Table 4 shows how this is possible, by illustrating the effects of variation in the parameters σ_z and κ that are held fixed in Table 3.

The rows of Table 4 correspond to alternative values of κ , while the columns correspond to alternative values of σ_z . In each cell corresponding to a particular pair (σ_z, κ) , I compute the value of θ required in order for the model to predict an average frequency of price changes of 0.184, as under the baseline calibration (the $\theta = 5$ row of Table 3), and then report the implied values of the statistics $E[|\Delta p|]$ and $K[|\Delta p|]$ as well. Thus the cells of the table represent different points on the

Figure 3 shows that the distribution of price changes is leptokurtic in those data as well.

³⁵For an alternative explanation of the observation of a leptokurtic distribution of individual price changes, see Midrigan (2008).

Table 4: Predictions regarding the size and frequency of price changes, for alternative values of σ_z and κ .

		$\sigma_z = 0.05$	$\sigma_z = 0.06$	$\sigma_z = 0.07$
$\kappa = 0.4$	θ	4.425	1.92	1.205
	$E(\Delta p)$	8.75	10.84	12.86
	$K(\Delta p)$	4.3	3.8	3.3
$\kappa = 0.5$	θ	12.475	5	2.8
	$E(\Delta p)$	8.64	10.51	12.47
	$K(\Delta p)$	4.8	4.1	3.8
$\kappa = 0.6$	θ	27.6	9.7	5.7
	$E(\Delta p)$	8.60	10.41	12.29
	$K(\Delta p)$	4.9	4.5	4.1

projection of a constant- $\bar{\Lambda}$ surface (within the three-dimensional space of possible model parameterizations) onto the $\sigma_z - \kappa$ plane. Our interest is in finding the point on this surface that also achieves certain target values for the other two statistics.

We observe from the table that in order for the model to predict an average size of price changes like that of the baseline calibration (between 10 and 11 percent) one needs a parameterization near the center column of the table, *i.e.*, a value of σ_z near 0.06. In order it to predict a coefficient of kurtosis only slightly above 4, one needs a parameterization close to the diagonal extending from upper left to lower right. Hence all three target values are achieved simultaneously at only one point in the three-dimensional space of parameterizations, corresponding to the center cell of the table (in boldface); these are the baseline parameter values used in the discussion above. In fact, the microeconomic data sets indicate an average size of price changes slightly larger than is implied by this parameterization (closer to 12 percent); arguably, a better fit with the micro evidence would be obtained by movement toward the lower right cell of Table 4 (meaning somewhat larger values of σ_z, κ , and θ). However, I have chosen the parameter values corresponding to the center cell as the baseline calibration, in order not to exaggerate the size of the information costs.

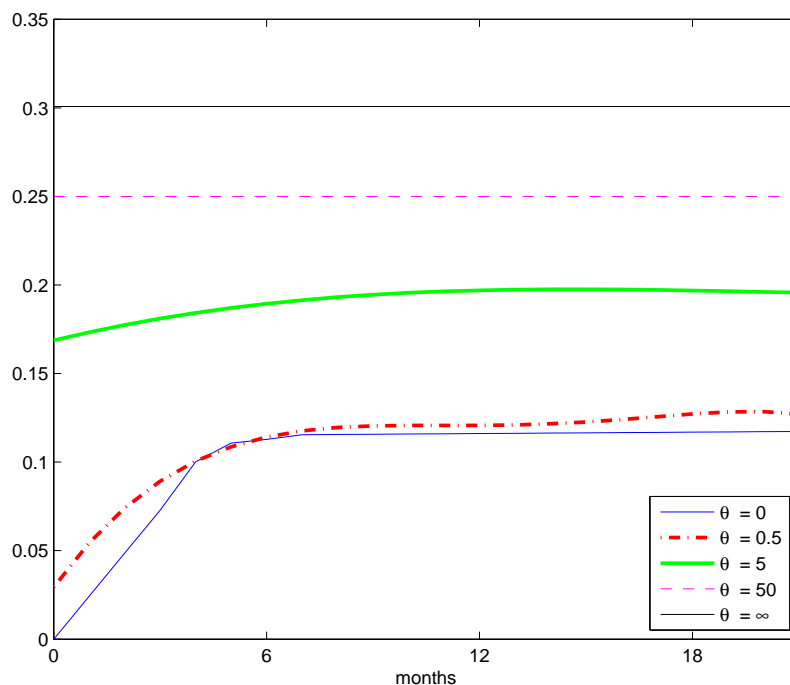


Figure 5: The probability of a price change as a function of time since the previous price change, for alternative values of θ .

3.4 Duration Dependence of Price Changes

Another feature of data on individual price changes that Klenow and Kryvtsov (2008) use to discriminate among alternative models of price adjustment is the hazard for price adjustment as a function of the time elapsed since the last price change. Figure 5 shows the duration dependence of the adjustment hazard in simulations of the stationary optimal price-review policy implied by the present model, for alternative values of θ (but again fixing all other parameters at their baseline values). In the case that $\theta = 0$ (the standard menu-cost model), there is essentially zero probability of another price change immediately following a price change (since $q_t(i)$ is equal to q^* immediately after any price change), while the probability of a price change is sharply rising with durations between zero and six months. For small enough positive values of θ (such as the value $\theta = 0.5$ shown in the figure), the hazard continues to be sharply rising with duration for the first few months. But for values of θ on the order of 5

or more, the probability of another price change in the first month following a price change is nearly as high as the probability in any of the later months; the hazard as a function of duration is nearly flat.

In fact, according to the statistics reported by Klenow and Kryvtsov for the BLS data from the top three urban areas (see their Figure 6), once one corrects for heterogeneity in the frequency of price changes in different types of goods, the hazard rate as a function of duration is very nearly flat over durations ranging from one to eleven months.³⁶ Klenow and Kryvtsov do report a substantial spike in the probability of a price change after exactly 12 months, but no increases in the hazard over the range of durations for which the hazard is steeply increasing for the models with low θ ;³⁷ nor is the increased hazard at 12 months an increase for all longer durations — the hazard is higher for durations of exactly 12 months than for either shorter or longer durations. This suggests possible time-dependence of the scheduling of price reviews, but is not consistent with a state-dependent pricing model with small information costs.³⁸

Thus we once again conclude that information costs on the order of $\theta = 5$ (if not even higher) are most consistent with the microeconomic evidence regarding the size and frequency of individual price changes. One observes (from Figure 2) that this size of information costs implies quite different dynamics of price adjustment than those of a standard (full-information) Ss model. Indeed, the hazard function is predicted to be quite flat, except in the case of quite large negative values of the normalized price (“price gaps” more negative than -.20). This suggests that the real effects of nominal disturbances under a realistic calibration may be closer to the predictions of the Calvo model than those of standard SDP models. I examine this question further in the next section.

³⁶Actually, they note a small *decline* in the hazard from the one-month duration to durations of two months or more.

³⁷Klenow and Kryvtsov count this as a problem for standard Ss models.

³⁸The fact that the spike is at precisely 12 months suggests that the reason is not that the likelihood of a price review depends on elapsed time since the last review in some sectors, but rather that in at least some sectors, there is significant seasonality of demand and/or costs; or alternatively, that it is less costly to condition the price review decision on the month of the year than on other sorts of information.

4 Monetary Non-Neutrality: A Simple Approach

What are the consequences of these results for the issue raised in the introduction, namely, whether endogenous timing of price reviews substantially reduces the real effects of monetary disturbances, relative to the prediction of an ET model with the same average frequency of price reviews? Because there are no aggregate shocks (and hence no monetary disturbances, among others) in the stationary equilibrium characterized in the previous section, the present model cannot, strictly speaking, give an answer to this question. Of course, the model could be extended, in a fairly straightforward way, to include aggregate nominal disturbances. But in this case, the distribution of normalized prices would no longer be constant over time, and the state space for each firm's decision problem would include a description of the current distribution \mathcal{Q}_t . For this reason, an exact model with aggregate disturbances would involve a state space of vastly higher dimension than in the case treated above, even if the dynamics of aggregate nominal expenditure are assumed to be extremely simple (for example, a random walk). This means that a proper treatment of the extension would require the introduction of approximation techniques different from any required in this paper, and such an analysis is deferred to another study.

Nonetheless, a simple exercise using the results computed above can give an indication of the degree to which the introduction of information costs changes the conclusions with regard to monetary non-neutrality relative to those obtained from a correspondingly parameterized full-information SDP model. Let us suppose that an exogenous shift in the log of aggregate nominal expenditure occurs, with the consequence that the normalized price of each firm i is shifted by the same amount, but that each firm's price-review decision is affected by this in the same way as in the stationary equilibrium with only idiosyncratic shocks. That is, I shall assume that each firm's price-review decision is based on a signalling mechanism that *would be optimal* under the assumption that aggregate disturbances never occur, though I am considering the consequences of an occurrence (at least once) of such a disturbance. While this involves a hypothesis of inattention that cannot be called fully "rational", the particular form of bounded rationality that is assumed may not be extremely implausible, given that under a realistic parameterization, typical aggregate nominal disturbances will be quite small relative to the idiosyncratic disturbances assumed in

the calibrated model above.³⁹

In fact, the simple exercise suffices to answer some questions about what one should expect to occur in the case of a genuinely constrained-optimal information structure. I consider the consequences of having each firm follow a policy that *would be optimal* in the case that the aggregate disturbances had no effect on the population distribution of normalized prices Q_t .⁴⁰ But in the case of *monetary neutrality*, as in the examples of Caplin and Spulber (1987) or Danziger (1999), this is true: an aggregate nominal disturbance has no effect on Q_t , even though many individual prices do not change in response to the shock. Hence *if* a similar neutrality result were to obtain in the extension of the present model to include aggregate shocks, the approximation proposed in this section would involve no inaccuracy. It follows that if the simple calculation shows that monetary neutrality does *not* obtain despite the endogeneity of the timing of price reviews, one can be certain that this would also be true of a model with a genuinely constrained-optimal information structure, though the *degree* of non-neutrality might well be different in the more sophisticated model.

Moreover, one can be fairly sure about the sign of the bias resulting from the simplifying assumption. In the case of a parameterization of the model of monopolistic competition that (as here) implies strategic complementarity among the pricing decisions of firms, the fact that firms are assumed to behave in a way that would be optimal if the nominal disturbance did not change the distribution of other firms' normalized prices — when in fact other firms' normalized prices are lowered (on average) by an increase in aggregate nominal expenditure — means that we are neglecting a reason why it should be optimal for firms to be *less inclined* to raise their prices in response to an increase in aggregate expenditure than is indicated by the simple calculation here. Hence the simple calculation (which, in essence, abstracts from the effects of strategic complementarity) is surely biased in the direction of indicating

³⁹Alternatively, the results presented in this section can be interpreted as optimal behavior under the assumption that each firm's objective is given by (1.1) with a period profit function given by (2.5), even if this form of objective can no longer be exactly justified by the microeconomic foundations proposed in section 2.

⁴⁰To be precise, the policy would be optimal if the parameter σ_z is set to reflect the standard deviation of innovations in the factor m_t due to *both* idiosyncratic and aggregate disturbances, and not just the innovations in the firm-specific factor $\log A_t(i)$. But since most of the variance of the innovations in m_t would be due to the idiosyncratic factor in any event, this correction to the above characterization of the optimal stationary policy would be a very small one.

more complete adjustment of prices (and hence smaller real effects) than would occur in the more sophisticated model, under a given parameterization.

Finally, while ignoring the way in which the dynamics of the aggregate distribution of normalized prices should affect the individual firm's decision problem introduces a bias into our estimate of the degree of monetary non-neutrality, it is not obvious that it should bias our conclusions about the degree to which the Calvo model (with its exogenous timing of price reviews) under-predicts the effective flexibility of prices. For abstracting from the consequences of strategic complementarity also biases the conclusions obtained with regard to monetary neutrality in the case of the Calvo model.⁴¹

Thus in the exercise considered here, I suppose that each firm's price-review policy continues to be described by the pair $\Pi \equiv (\Lambda, q^*)$ that represent a stationary optimal policy in the case that only idiosyncratic disturbances exist, but I replace the law of motion (1.3) with

$$q_{t+1}(i) = q_t(i) + z_{t+1}(i) - \nu_{t+1},$$

where $z_{t+1}(i)$ is drawn independently for each firm from the distribution $g(z)$ as before, but the additional term ν_{t+1} represents an unexpected permanent change in $\log P_{t+1}Y_{t+1}$ relative to the value of $\log P_tY_t$. The question that we wish to address is the extent to which such an aggregate shock changes the average level of prices as opposed to aggregate real activity.

A quantity of interest is therefore

$$h(\nu) \equiv E_i[\Delta p_{t+1}(i) | \nu_{t+1} = \nu],$$

the average price increase resulting from an innovation of size ν in aggregate nominal expenditure. I assume that the population distribution of values of $q_{t+1}(i)$ prior to the aggregate shock is given by the stationary distribution \bar{Q} computed above for the case in which only idiosyncratic shocks exist.

There are two simple benchmarks with which it is useful to compare the function $h(\nu)$ obtained for the model with information-constrained price review decisions. One is the benchmark of *perfect neutrality*. In this case (represented, for example, by the

⁴¹The results presented below for the "Calvo model" refer to a model in which the timing of price reviews is exogenous, but in which firms that review their prices seek to maximize the approximate objective described by (1.1) and (2.5), rather than the correct objective for the model of monopolistic competition with aggregate disturbances.

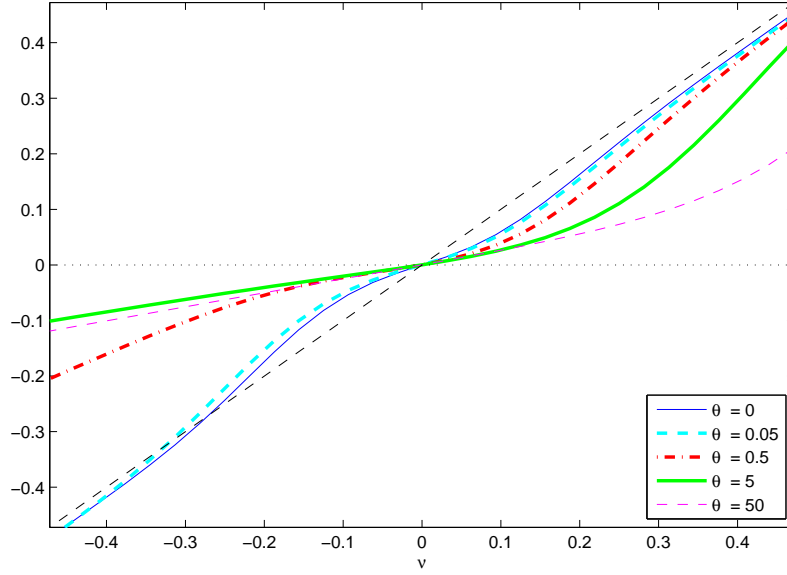


Figure 6: The function $h(\nu)$, for alternative values of θ . The dashed line on the diagonal shows the benchmark of perfect neutrality.

SDP models of Caplin and Spulber or Danziger), $h(\nu) = \nu$, a straight line with a slope of 1. Another useful benchmark is the prediction of the *Calvo model* of price adjustment, when calibrated so as to imply an average frequency of price change equal to the one that is actually observed, $\bar{\Lambda}$. In this case, $h(\nu) = \bar{\Lambda}\nu$, a straight line with a slope $\bar{\Lambda} < 1$.⁴²

Figure 6 plots the function $h(\nu)$, for each of the several possible values of θ considered in Table 1. The figure also plots the benchmark of full neutrality (shown as a dashed line on the diagonal).⁴³ One observes that in all cases, there is less than full immediate adjustment of prices to a purely monetary shock, in the case of small shocks ($0 < h(\nu) < \nu$ for small $\nu > 0$, and similarly $\nu < h(\nu) < 0$ for small $\nu < 0$).

⁴²For each firm i that reviews its price in period $t + 1$, the log price change is equal to $d_{t+1}(i) \equiv q^* - q_t(i) - z_{t+1}(i) + \nu_{t+1}$, where here $q_t(i)$ means the firm's normalized price *after* any period t preview. In the Calvo model, a fraction $\bar{\Lambda}$ of the firms review their prices, and these represent a uniform sample from the population of firms, so that the mean log price change is $\bar{\Lambda}E_id_{t+1}(i)$. The Calvo model further implies that $E_iq_t(i) = q^*$, so that $E_id_{t+1}(i) = \nu_{t+1}$.

⁴³The Calvo benchmark cannot be plotted as any single line in this figure, as it depends on the value of $\bar{\Lambda}$, and the value of $\bar{\Lambda}$ is different for the different values of θ , as shown in Table 5.

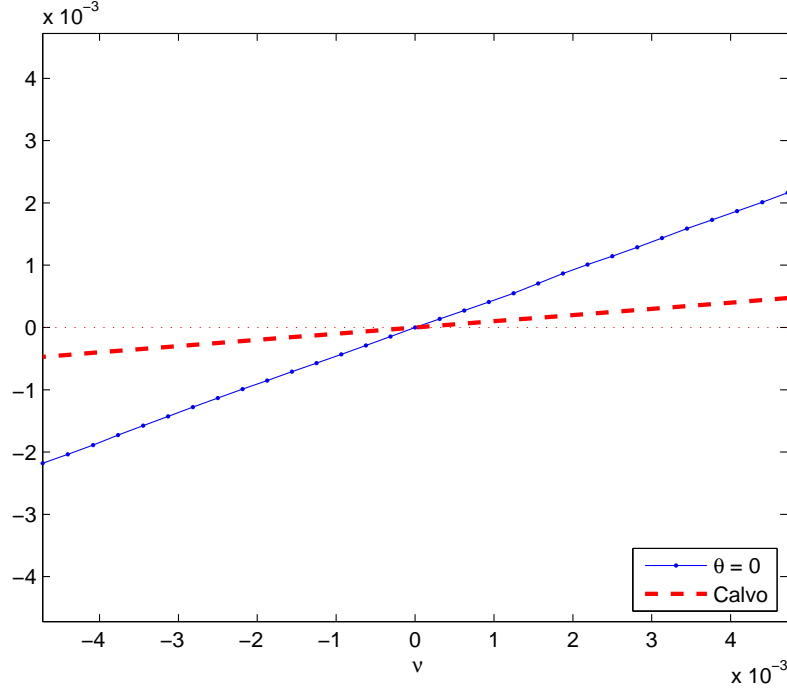


Figure 7: A closer view of the function $h(\nu)$, for the case $\theta = 0$. The dashed line shows the prediction of the Calvo model for purposes of comparison.

However, there is greater proportional adjustment to larger shocks, and in fact in each case the graph of $h(\nu)$ eventually approaches the diagonal (the benchmark of full neutrality) for large enough positive shocks. The size of shocks required for this to occur, though, is greater the larger is θ . There is also an evident asymmetry in the responses to large shocks, in the case of finite positive values of θ ; there need not be a full immediate adjustment of the average price to the nominal disturbance even in the case of very large negative shocks.⁴⁴ This is because in the model of monopolistic competition proposed above, losses are bounded no matter how much too high one's price may be, whereas unbounded losses are possible in the case of a price that is too low.

Even in the case of shocks of a magnitude that does not result in full adjustment

⁴⁴Of course, it is hard to be sure how much weight to attach to this result, given that the approximation involved in neglecting the effects of the aggregate shock on the population distribution of normalized prices (when deriving the price-review policies of the individual firms) is particularly unappealing in the case of a very large negative shock.

in either case, a positive shock results in more nearly complete price adjustment, on average, than does a negative shock of the same size. This is a direct result of the asymmetry of the optimal hazard function, already noted in Figure 3. (Because firms with prices that are too low are more likely to immediately adjust their prices than firms with prices that are too high, more adjustment occurs immediately in response to a positive shock than to a negative shock.) The result implies, in turn, that the effects of a contraction of nominal aggregate demand on real activity will be greater than the effects of an expansion of nominal aggregate demand by the same number of percentage points; for more of the positive demand disturbance will be dissipated in an immediate increase in prices than occurs in the case of a negative disturbance. This conclusion, of course, echoes a feature often found in old-fashioned Keynesian models, which assumed that prices (or wages) were “downwardly rigid” but not upwardly rigid to the same extent. The present model justifies similar behavior as a consequence of optimization; but the reason here is not any resistance to price declines — instead, firms are *more worried* about allowing their prices to remain too low than they are about allowing them to remain too high.

Even in the case of small shocks, while there is not full adjustment to monetary shocks in the month of the shock, the average price increase is *many times larger* than would be predicted by the Calvo model, in the case of sufficiently small values of θ . Figure 7 shows a magnified view of the graph of $h(\nu)$ for small values of ν , in the case $\theta = 0$, with the prediction of the Calvo model also shown by a dashed line. The slope of the curve $h(\nu)$ near the origin is several times greater than $\bar{\Lambda}$, the slope predicted by the Calvo model. This confirms the finding of authors such as Golosov and Lucas (2007) about the consequences of SDP under full information.

However, for larger values of θ , the Calvo model provides quite a good approximation, in the case of small enough shocks. Figure 8 shows a similarly magnified view of the graph of $h(\nu)$ in the case $\theta = 5$. One observes that the prediction of the Calvo model is quite accurate,⁴⁵ for shocks of the magnitude shown in the figure. In fact, these shocks (up to half a percent innovation in the long-run forecast of the price level, in a single month) are quite large relative to typical nominal disturbances in an economy like that of the US. So while the Calvo model is much less accurate in

⁴⁵Note that this is a differently parameterized Calvo model than in Figure 7: the frequency of price review in the Calvo model is adjusted to match the higher average frequency of price reviews $\bar{\Lambda}$ in the model with $\theta = 5$.

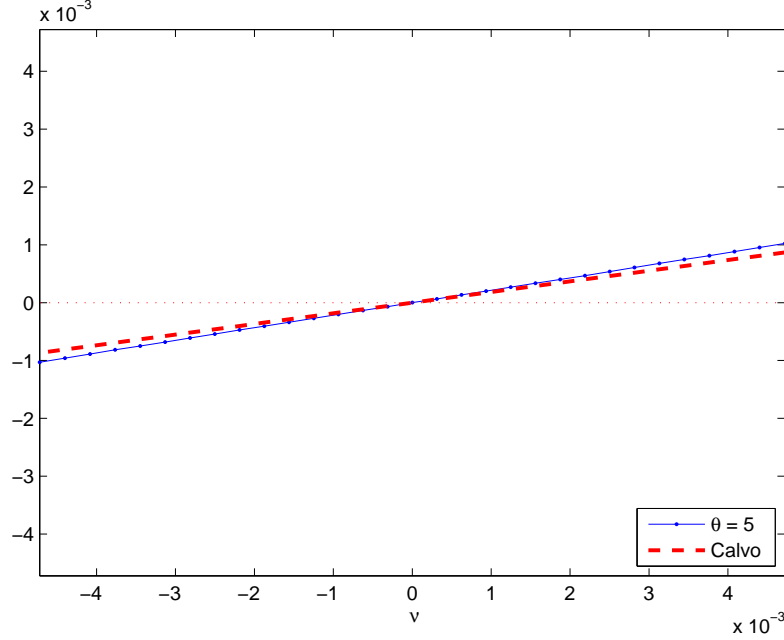


Figure 8: A closer view of the function $h(\nu)$, for the case $\theta = 5$. The dashed line again shows the prediction of the Calvo model.

its predictions about large nominal disturbances (especially, large positive shocks), as shown in Figure 6, in the case that $\theta = 5$ one should expect its predictions to be reasonably accurate in the case of most of the shocks that occur with any frequency. For even larger values of θ , the approximation is even better, and the range over which the approximation is accurate extends to even larger shock sizes.

One way of measuring the extent to which the inaccuracy of the Calvo approximation matters in general is by considering the slope of a linear regression of the log price change on the size of the current aggregate shock. Suppose that aggregate nominal disturbances ν_{t+1} occurs in each of a large number of periods, drawn independently each time from a distribution $N(0, \sigma_\nu^2)$, and that each time price-review decisions are made in the way assumed above. Suppose that we collect data on individual price changes in each of these periods, and then approximate the function $h(\nu)$ by a linear equation,

$$\Delta p_{t+1}(i) = \alpha + \beta \nu_{t+1} + \epsilon_{t+1}(i),$$

where the residual is assumed to have mean zero and to be orthogonal to the aggregate

Table 5: The coefficient β from a regression of log price changes on the current monetary shock, for alternative values of θ . The value of $\bar{\Lambda}$ implied by the stationary optimal policy in each case is shown for purposes of comparison. (Both quantities reported in percentage points.)

θ	$\bar{\Lambda}$	β
0	10.0	45.3
.05	9.0	40.3
.5	10.4	25.2
5	18.4	22.0
50	25.0	25.5
∞	30.1	30.1

shock.⁴⁶

The values of β obtained from simulations of the stationary optimal policies corresponding to the different values of θ are given in Table 5, which also reports the values of $\bar{\Lambda}$ implied by each of these policies. (In these simulations I use the value $\sigma_\nu = .001$. This corresponds to a standard deviation for quarterly innovations in the long-run price level of approximately 17 basis points.) One observes that the Calvo model under-predicts the flexibility of prices very substantially in the full-information case ($\theta = 0$), which is to say, in a standard SDP model of the kind studied by Golosov and Lucas (2007). For the parameter values assumed here, I find that the correct linear response coefficient is more than 4.5 times as large as the one predicted by the Calvo model. But in the baseline case ($\theta = 5$), the correct coefficient is only 20 percent larger than the prediction of the Calvo model. The Calvo model is even more accurate if information costs are larger; for example, if $\theta = 50$, it under-predicts the immediate price response by only 2 percent. In the limiting case of unboundedly large θ , the Calvo model is perfectly accurate.

Further work will be required to determine the degree to which similar conclusions obtain if one allows firms to be informed to an optimal extent about both idiosyncratic and aggregate disturbances, taking into account the *ex ante* joint distribution of both

⁴⁶Here I imagine that and estimate the coefficients α and β by ordinary least squares. Under the full neutrality benchmark, β would equal 1; the Calvo model predicts that β should equal $\bar{\Lambda}$.

kinds of disturbances. However, the convergence of the optimal hazard function to a constant function (*i.e.*, the prediction of an appropriately calibrated Calvo model) as θ is made unboundedly large should still occur, even in the case of a much larger state space, for the same reason that it occurs here in the model with only idiosyncratic shocks. Hence one expects to find, even in a more sophisticated model, that the Calvo model becomes a good approximation in the case of high enough values of θ (especially for small aggregate disturbances); the only question is the exact rate at which this convergence occurs.

5 Conclusion

I have presented a model in the timing of price changes results from optimizing behavior on the part of firms subject to a fixed cost of conducting a review of existing pricing policy. Standard models of state-dependent pricing, however, are generalized by assuming that a firm’s policy with regard to the timing of price reviews is designed to economize on the cost of being continuously informed about market conditions during the intervals between full-scale reviews. The introduction of interim information costs softens the distinction, emphasized in prior contributions, between the dynamics of price adjustment in models with exogenous timing of price adjustments and models with state-dependent pricing, by attenuating both the “selection effect” emphasized by Golosov and Lucas (2007) and the relative importance of the “extensive margin of price adjustment” emphasized by Caballero and Engel (2007). In the limiting case of sufficiently large interim information costs, the predicted dynamics of price adjustment are identical to those of the Calvo (1983) model of staggered price-setting.

At a minimum, this result means that there is no reason to regard the predictions of (full-information) “menu cost” models as more likely to be accurate than the predictions of the Calvo model, simply on the ground that the former models have firmer foundations in optimizing behavior. Both models appear as nested (extreme) cases of the more general model presented here, so that the question of which special case is more reliable as an approximation is a quantitative matter, rather than one that can be settled simply on the basis of the appeal of optimizing models.

The illustrative calculations presented in section 4 furthermore suggest that a model with interim information costs of moderate size may imply aggregate behavior fairly similar to that predicted by the Calvo model, and quite different from that pre-

dicted by a full-information menu-cost model. Further work is needed to investigate to what extent this conclusion obtains when the endogenous information structure takes account of the *ex ante* possibility of aggregate as well as idiosyncratic shocks. But these calculations show that it is possible for predictions of the Calvo model to be fairly accurate for many purposes — predicting the aggregate responses to disturbances of the magnitude that occur at most times — in a model that does not possess certain features of the Calvo model that are often argued to be implausible. In particular, the model with a finite positive interim cost of information does *not* imply that prices are equally unlikely to be adjusted even when a given firm’s price happens over time to have become far out of line with profit maximization, or even when very large disturbances affect the economy. However, because firms are in these situations only very infrequently, the predictions of the Calvo model may nonetheless be relatively accurate much of the time.

It is important to note, however, that the implications of the present model are likely to differ from those of the Calvo model in important respects, even if a relatively large value of θ is judged to be empirically realistic. First, even if the price adjustments predicted by this model are similar to those of the Calvo model under all but extreme circumstances, the model’s predictions under extreme circumstances may be of disproportionate importance for calculations of the welfare consequences of alternative stabilization policies, as argued by Kiley (2002) and Paustian (2005). And second, even in the limit of an unboundedly large value of θ (so that no interim information is available at all), the present model’s predictions differ from those of the Calvo model in at least one important respect: the equilibrium frequency of price review $\bar{\Lambda}$ is endogenously determined, rather than being given exogenously. In particular, the value of $\bar{\Lambda}$ is unlikely to be policy-invariant; for example, one would expect it to be higher in the case of a higher average inflation rate, as in the generalized Calvo model of Levin and Yun (2007). For this reason as well, the present model may well have different implications than the Calvo model for the welfare ranking of alternative policy rules, as in the analysis of Levin and Yun. This is another important topic for further study.

References

- [1] Bilal, Mark, and Peter J. Klenow, “Some Evidence on the Importance of Sticky Prices,” *Journal of Political Economy* 112: 947-985 (2004).
- [2] Bonomo, Marco, and Carlos Carvalho, “Endogenous Time-Dependent Rules and Inflation Inertia,” *Journal of Money, Credit and Banking* 36: 1015-1041 (2004).
- [3] Burstein, Ariel T., “Inflation and Output Dynamics with State Dependent Pricing Decisions,” *Journal of Monetary Economics* 53: 1235-1257 (2006).
- [4] Burstein, Ariel T., and Christian Hellwig, “Prices and Market Share in a Menu Cost Model,” unpublished, U.C.L.A., September 2007.
- [5] Caballero, Ricardo J., “Time Dependent Rules, Aggregate Stickiness, and Information Externalities,” Discussion Paper no. 428, Department of Economics, Columbia University, May 1989.
- [6] Caballero, Ricardo J., and Eduardo Engel, “Microeconomic Rigidities and Aggregate Price Dynamics,” *European Economic Review* 37: 697-717 (1993a).
- [7] Caballero, Ricardo J., and Eduardo Engel, “Microeconomic Adjustment Hazards and Aggregate Dynamics,” *Quarterly Journal of Economics* 108: 359-383 (1993b).
- [8] Caballero, Ricardo J., and Eduardo Engel, “Explaining Investment Dynamics in U.S. Manufacturing: A Generalized (S,s) Approach,” *Econometrica* 67: 783-826 (1999).
- [9] Caballero, Ricardo J., and Eduardo Engel, “Price Stickiness in Ss Models: Basic Properties,” unpublished, M.I.T., October 2006.
- [10] Caballero, Ricardo J., and Eduardo Engel, “Price Stickiness in Ss Models: New Interpretations of Old Results,” *Journal of Monetary Economics* 54 (supp. 1): 100-121 (2007).
- [11] Calvo, Guillermo A., “Staggered Prices in a Utility Maximizing Framework,” *Journal of Monetary Economics* 12: 383-398 (1983).

- [12] Caplin, Andrew S., and Daniel F. Spulber, "Menu Costs and the Neutrality of Money," *Quarterly Journal of Economics* 102: 703-726 (1987).
- [13] Cover, Thomas M., and Joy A. Thomas, *Elements of Information Theory*, New York: Wiley-Interscience, 2d ed., 2006.
- [14] Danziger, Leif, "A Dynamic Economy with Costly Price Adjustments," *American Economic Review* 89: 878-901 (1999).
- [15] Devereux, Michael B., and James Yetman, "Predetermined Prices and the Persistent Effects of Money on Output," *Journal of Money, Credit and Banking* 35: 729-742 (2003).
- [16] Dotsey, Michael and Robert G. King, "Implications of State-Dependent Pricing for Dynamic Macroeconomic Models," *Journal of Monetary Economics* 52: 213-242 (2005).
- [17] Dotsey, Michael, Robert G. King, and Alexander L. Wolman, "State Dependent Pricing and the General Equilibrium Dynamics of Money and Output," *Quarterly Journal of Economics* 114: 655-690 (1999).
- [18] Eichenbaum, Martin, Nir Jaimovich, and Sergio Rebelo, "Reference Prices and Nominal Rigidities," unpublished, Northwestern University, March 2008.
- [19] Gertler, Mark, and John Leahy, "A Phillips Curve with an Ss Foundation," unpublished, New York University, May 2007.
- [20] Golosov, Michael, and Robert E. Lucas, Jr., "Menu Costs and Phillips Curves," *Journal of Political Economy* 115: 171-199 (2007).
- [21] Kiley, Michael T., "Partial Adjustment and Staggered Price Setting," *Journal of Money, Credit and Banking* 34: 283-298 (2002).
- [22] Klenow, Peter J., and Oleksiy Kryvtsov, "State-Dependent or Time-Dependent Pricing: Does it Matter for Recent U.S. Inflation?" *Quarterly Journal of Economics* 123: 863-904 (2008).
- [23] Levin, Andrew, and Tack Yun, "Reconsidering the Natural Rate Hypothesis in a New Keynesian Framework," *Journal of Monetary Economics* 54: 1344-1365 (2007).

- [24] Matejka, Filip, “Rationally Inattentive Seller: Sales and Discrete Pricing,” unpublished, Princeton University, August 2008.
- [25] Midrigan, Virgiliu, “Menu Costs, Multi-Product Firms, and Aggregate Fluctuations,” unpublished, New York University, August 2008.
- [26] Miranda, Mario J., and Paul L. Fackler, *Applied Computational Economics and Finance*, Cambridge: M.I.T. Press, 2002.
- [27] Nakamura, Emi, and Jon Steinsson, “Five Facts about Prices: A Reevaluation of Menu Cost Models,” *Quarterly Journal of Economics* 123: 1415-1464 (2008a).
- [28] Nakamura, Emi, and Jon Steinsson, “Monetary Non-Neutrality in a Multi-Sector Menu Cost Model,” unpublished, Columbia University, April 2008b.
- [29] Paustian, Matthias, “The Role of Contracting Schemes for the Welfare Costs of Nominal Rigidities over the Business Cycle,” unpublished, University of Bonn, April 2005.
- [30] Phelps, Edmund S., *Seven Schools of Macroeconomic Thought*, Oxford: Clarendon Press, 1990.
- [31] Reis, Ricardo, “Inattentive Producers,” *Review of Economic Studies* 73: 793-821 (2006).
- [32] Rubinstein, Ariel, “Finite Automata Play the Repeated Prisoner’s Dilemma,” *Journal of Economic Theory* 39: 83-96 (1986).
- [33] Shannon, Claude E., “A Mathematical Theory of Communication,” *Bell System Technical Journal* 27: 379-423 and 623-656 (1948).
- [34] Sims, Christopher A., “Stickiness,” *Carnegie-Rochester Conference Series on Public Policy* 49: 317-356 (1998).
- [35] Sims, Christopher A., “Implications of Rational Inattention,” *Journal of Monetary Economics* 50: 665-690 (2003).
- [36] Sims, Christopher A., “Rational Inattention: A Research Agenda,” unpublished, Princeton University, March 2006.

- [37] Woodford, Michael, *Interest and Prices: Foundations of a Theory of Monetary Policy*, Princeton: Princeton University Press, 2003.
- [38] Woodford, Michael, “Inattention as a Source of Randomized Discrete Adjustment,” unpublished, Columbia University, April 2008.
- [39] Zbaracki, Mark, Mark Ritson, Daniel Levy, Shantanu Dutta, and Mark Bergen, “Managerial and Customer Costs of Price Adjustment: Direct Evidence from Industrial Markets,” *Review of Economics and Statistics* 86: 514-533 (2004).