Review of Economic Studies (2006) **73**, 793–821 0034-6527/06/00290793$02.00

Inattentive Producers

RICARDO REIS
Princeton University, NBER, and CEPR

First version received June 2004; *final version accepted November* 2005 (*Eds.*)

I present and solve the problem of a producer who faces costs of acquiring, absorbing, and processing information. I establish a series of theoretical results describing the producer's behaviour. First, I find the conditions under which the producer prefers to set a plan for the price he or she charges, or instead prefers to set a plan for the quantity he or she sells. Second, I show that the agent rationally chooses to be inattentive to news, only sporadically updating his or her information. I solve for the optimal length of inattentiveness and characterize its determinants. Third, I explicitly aggregate the behaviour of many such producers. I apply these results to a model of inflation. I find that the model can fit the quantitative facts on post-war inflation remarkably well, that it is a good forecaster of future inflation, and that it survives the Lucas critique by fitting also the pre-war facts on inflation moderately well.

1. INTRODUCTION

A long-standing question in macroeconomics is why don't prices adjust every instant to reflect the incoming stream of news on the environment facing firms? This question is important because its answer determines the answer to many other questions in macroeconomics. For instance, the imperfect adjustment of prices to news on money lies behind the effects of monetary policy on real activity. To give another example, if we can understand the dynamic response of prices to shocks, we should be able to explain the dynamics of inflation, one of the key aggregate variables that macroeconomists purport to explain.

At least since John Maynard Keynes, a popular answer has been to assume that prices are fixed for periods of time. Barro (1972), Sheshinski and Weiss (1977, 1982), Rotemberg (1982), and Mankiw (1985) provided a micro-foundation for sticky prices by assuming that there is a fixed physical cost that firms must pay whenever they change their price. Caballero and Engel (1991) and Caplin and Leahy (1997) aggregated this infrequent adjustment across many different firms. Danziger (1999), Dotsey, King and Wolman (1999), and Golosov and Lucas (2003) studied the effects of monetary policy in these economies. A closely related model of sticky prices bypasses the micro-foundations and assumes from the start that prices adjust only at some random dates picked from a specific distribution that allows for simple aggregation (Calvo, 1983; Woodford, 2003*a*).

The model of sticky prices has always been criticized, but over the past decade the criticism has intensified. Researchers have noted that there is little support in the data for the model's basic assumption. With the exception of magazine prices and restaurant menus, for most products it is difficult to identify any significant fixed physical costs of changing prices. Research has also found that the data do not support the model's key micro-prediction. Bils and Klenow (2004) noted that individual prices change very frequently in the U.S. Finally, many authors (*e.g.* Mankiw, 2001) have shown that the macroeconomic predictions of the sticky-price model for the relation between inflation, real activity, and monetary policy are counterfactual.

An alternative explanation for the imperfect adjustment of prices to news acknowledges that people have limited information and a limited ability to perform computations. These models start by emphasizing that in the standard classical model, agents are aware of all the information every instant and are constantly using it to compute their optimal actions. Yet, there is an enormous amount of information in the world and most of it comes with a cost, in money or time, both in acquiring the information but especially in interpreting it. The limited-information approach argues that following the hallmark of economics of studying choice subject to constraints, information should be treated as a costly good. The Lucas (1972) islands model showed that if price-setters have imperfect information, they will adjust incompletely to news, which generates nominal rigidities and real effects of monetary policy. More recently, Mankiw and Reis (2002) provided a limited-information alternative to the sticky-price Calvo (1983) model by assuming that agents update their information sets and price plans at randomly chosen dates. They showed that this model of sticky information is able to match some facts on inflation and output dynamics and to generate reasonable responses of these variables to monetary policy shocks.

Currently though, models of pricing based on limited information lack a micro-foundation based on optimizing behaviour, lack an explicit aggregation across many agents, and lack an explicit contrast of their predictions with the data on inflation. This is what this paper proposes to do. I will use the inattentiveness model of limited information to model the behaviour of producers. This model adds to a standard profit-maximization problem, one new constraint: that agents must pay a cost to acquire, absorb, and process information in forming expectations and making decisions. The basic implication of this assumption is that agents rationally choose to be inattentive, only sporadically updating their information sets and plans at optimally chosen dates.

The model makes several interesting novel predictions. First, it provides a micro-foundation for time-contingent adjustment, since people update their plans at certain dates regardless of the state of the economy at these dates. Time-contingent adjustment is appealing relative to its state-contingent alternative because it typically implies larger and longer real effects of monetary policy and it can reproduce the delayed and hump-shaped response of inflation to shocks. Second, I derive the conditions under which producers choose to write plans for the price they will charge or for the quantity they will produce. These conditions are easy to verify empirically and they provide a rich set of theoretical predictions that can be used both to test the model as well as to build future models of inflation in which industries differ over their type of plan. Third, I derive an approximate solution for the producer's decision of how long to be inattentive for. One virtue of this solution is that it shows how industry characteristics affect the frequency of planning. Fourth, I explicitly aggregate the behaviour of many inattentive producers. One surprising result emerges: under some general conditions, the distribution of inattentiveness is exponential. Fifth, under some particular assumptions, these theoretical results turn out to provide a micro-foundation to the assumptions made by Mankiw and Reis (2002), thus putting in firm ground this particular model of nominal rigidities due to incomplete information. Finally and sixth, this paper exploits these micro-foundations to construct a simple model of inflation and show that it performs well in fitting the data in three aspects: the model can quantitatively match the second moments of the post-war data very closely, it beats reduced-form autoregressive models at forecasting inflation, and it is robust to the Lucas critique in that it can account moderately well for the inflation data under a different policy regime in the pre-war U.S.

There are a few papers that are more closely related to this one. Caballero's (1989) derivation of time-dependent rules from first principles is a precursor to some of the calculations in this paper. He considers a more restricted choice of planning dates though, and focuses on a different set of issues. Bonomo and Carvalho (2004) provide a model of optimal time-contingent price adjustment, but one in which prices must be fixed in between adjustments rather than following possibly time-varying plans, as in the model in this paper. Burstein (2006) presents a sticky-plan model in which prices also follow predetermined plans that are only sporadically updated. The price-setters in his model have full information each instant and use it to decide whether to adjust their plan. They choose to respond differently to large and small shocks, and to positive and negative desired price changes. In the model in this paper instead, consistent with the underlying assumption that information is costly, not just price plans but also information sets are updated sporadically. Because they are not aware of the news as it arrives, price-setters do not respond asymmetrically to it.

Finally, Woodford (2003*b*) and Moscarini (2004) model inattentiveness by price-setters using an alternative approach suggested by Sims (2003). Using tools from electrical engineering, this approach models agents who have a limited capacity to absorb information. In Moscarini's (2004) version, agents also choose to only infrequently update their information, and when they do, they only obtain an imperfect signal on the state of the world. In the inattentiveness model instead, when producers update their plans, they become aware of everything that is relevant. While this feature of the inattentiveness model is extreme, it has the pay-off of making the model significantly more tractable. For instance, the theoretical results in this paper have already led to fully specified models of inflation, whereas theories of limited information capacity still face some difficult (but exciting) conceptual hurdles. Awaiting progress in those theories, this paper's development of the inattentiveness model should make possible future comparisons of the two approaches.

The paper is organized as follows. Section 2 states the problem facing producers. Section 3 answers a first question: will the inattentive agent set a plan for prices or for quantities? Section 4 solves the problem of how often to adjust and examines the determinants of inattentiveness. Section 5 aggregates the behaviour of many inattentive agents. Section 6 uses these results to set up a model of inflation. Section 7 contrasts the model with data, and Section 8 concludes.

2. THE INATTENTIVE PRODUCER'S PROBLEM

2.1. *An informal description of the problem*

Consider the problem facing a monopolist that produces a perishable good. Both the production technology and the demand for its good are uncertain and can change every instant, so that to obtain the full information first-best profits, the producer would have to observe the determinants of costs and demand every instant. The assumption in this paper is that this entails a cost, namely, that it is costly to acquire, absorb, and process information. It is costly to acquire information in the sense of collecting all the pieces of information that are relevant to assess the current state of the world. It is costly to absorb information in the sense of compiling this information into the relevant sufficient statistics needed to make optimal decisions. And it is costly to process information in the sense of coming up with the optimal action and implementing it.

For a typical producer, these costs stand, for instance, for the costs of keeping detailed accounts of sales, the costs of monitoring and assessing the different stages of production, and the payments to outside consultants for their advice. Radner (1992) insightfully observed that a large fraction of the workforce is employed in managerial occupations, which are essentially about processing information and making decisions. Even if only a small fraction of these people's time is spent at acquiring, absorbing, and processing information towards making optimal decisions, the costs of doing so can be substantial. Zbaracki, Ritson, Levy, Dutta and Bergen (2004) directly measured the costs incurred by a large U.S. manufacturing firm associated with setting its price catalogue. These were as high as 1·2% of the company's revenue and 20% of its net margin.

Facing these costs, the producer optimally chooses to only update his or her information sporadically, and to be inattentive to new information in between adjustment dates. When the producer does obtain information, conditional on it he or she decides whether to set prices or quantities, which price to charge or which quantity to sell for the duration of the plan, and when next to plan.

To illustrate these three simultaneous decisions, consider the example of a fictional baker. His or her first decision is on which variable to write a plan on: price or quantity. If the baker sets a price for the bread, he or she will keep the oven burning and bread coming out as long as customers are walking through the door. If instead the baker chooses to produce a certain amount of bread, he or she then gives it to a seller. This seller takes the bread to the market and distributes it among homes and shops, charging whatever positive price is necessary to sell all the bread today, since by the end of the day the bread becomes stale and worthless. Finally, the seller returns the sales proceeds to the baker. Facing this choice of prices vs. quantities, the baker forms an expectation of his or her profits under the two alternatives and chooses the most profitable one.

In both cases, an important assumption that I maintain is market clearing. In the case of a price plan, this implies that some mechanism in the economy directs consumers to the baker's shop as long as their marginal utility of bread is above the posted price. In the case of a quantity plan, there is some mechanism in the economy that acts as a seller finding the price that clears the market. These mechanisms serve the purpose of the fictional Walrasian auctioneer that economists routinely assume to ensure that markets clear in equilibrium. I maintain this assumption because an operational non-market-clearing definition of equilibrium is still an elusive research objective, despite the initial steps of Barro and Grossman (1971).

The second decision for the baker is on the content of the plan. If the baker chooses a price plan, this consists of the path of prices to charge until the next planning date. If he or she chooses a quantity plan, it is the path for the amount of bread to produce. In both cases, note that the producer chooses paths, not numbers, since he or she knows the calendar date and so the baker's conduct can depend on it. Recent empirical research on the dynamics of individual prices supports this path setting. Many prices are revised at infrequent intervals, but in between revisions have large predetermined swings that coincide with predictable events, such as holidays, the Christmas shopping period, or weeks of "specials" (Chevalier, Kashyap and Rossi, 2003; Rotemberg, 2005).

The final decision at a planning date is on the horizon of the plan, or on when to obtain new information and plan again. The baker realizes that while on the one hand, extending the horizon of the plan saves on the costs of planning, on the other hand, it implies that decisions towards the end of the plan are made with severely outdated information and so are likely to be missing on substantial profits. Sufficiently far in the future, the cost of following an outdated plan becomes too high relative to the cost of obtaining information, and it is optimal to stop and plan again.

Readers might wonder whether there isn't some information that the producer can obtain for free. For instance, why can't the baker observe the quantity sold at the end of the day at his or her fixed price, or hear from the seller at which price did the seller sell the bread? The answer is that, in principle, the baker can. But then, the baker must use this one piece of information to infer the current state of demand, and proceed to collect a myriad of other pieces of information that affect demand in the future, from consumers' taste for bread to their disposable income. Moreover, the baker must go through the entire production process and realize how much exactly he or she paid for each input and how long it took to combine them to make bread, as well as forecast how all of these are expected to change by tomorrow. Even if some of the information is costless to acquire, it is still costly to absorb and process this information to change the optimal plan. The basic assumption in this paper is not inconsistent with people being aware of some events, as long as it is still costly to think through this information.[1] Moreover, as I will show later, even tiny costs of information can generate substantial inattentiveness.

1. An alternative assumption is that the producer can acquire a few pieces of information every instant, absorb these into a sufficient statistic, and use these to evaluate an optimal plan, all at no cost. Still, as long as there is some other independent information that can only be acquired, absorbed, and processed at a cost, the model in this paper is still applicable. The inattentiveness is now only with respect to the costly pieces of information.

2.2. *The formal problem*

The monopolist produces a single perishable good with a stochastic technology represented by a continuous and smooth cost function $C(Y, \boldsymbol{s}) : \mathbb{R}^{S+1} \to \mathbb{R}$. The quantity produced is denoted by Y and $\boldsymbol{s}$ is a vector stochastic process with S components standing for the different relevant bits of information. The demand for this product is also stochastic and is represented by the continuous and smooth function $Q(P, \boldsymbol{s}) : \mathbb{R}^{S+1} \to \mathbb{R}$, where P stands for the price charged. I assume that demand is always positive and falls with the price being charged.

The stochastic process $\boldsymbol{s}_t$ is defined on a standard filtered probability space with filtration $F = \{F_t, t \geq 0\}$. I assume that $\boldsymbol{s}_t$ has the Markov property and, without loss of generality, that it is arranged so that it is first-order Markov. The state at a given date $t + \tau$ is then a function of $\boldsymbol{s}_t$ and a set of innovations $u^\tau = (u_t, u_{t+\tau}]$, so that I can write $\boldsymbol{s}_{t+\tau} = \Psi(\boldsymbol{s}_t, u^\tau)$ to denote the transition between the state at date t and the state at date $t + \tau$, which is assumed to be differentiable.

The planning dates are denoted by the almost surely non-decreasing function $D(i) : \mathbb{N}_0 \to \mathbb{R}$ with $D(0) = 0$. The periods of inattentiveness are defined as $d(i) = D(i) - D(i-1)$. The optimal choice of planning dates defines a new filtration $\Im = \{\Im_t, t \geq 0\}$ such that $\Im_t = F_{D(i)}$ for $t \in [D(i), D(i+1))$. The restriction imposed by a plan is that the producer's choices at time t must be measurable with respect to $\Im$. That is, the producer's choices for time t must be conditional on the information he or she has at time t, which coincides with the available information in the economy at the last planning date.

The producer maximizes expected profits conditional on his or her information. If at time t the producer sets a price, he or she obtains profits:[2]

$$\Pi^P(\boldsymbol{s}_{D(i)}, t - D(i)) = \max_{P_t} E[P_t Q(P_t, \boldsymbol{s}_t) - C(Q(P_t, \boldsymbol{s}_t), \boldsymbol{s}_t) \mid \Im_t]. \tag{1}$$

The solution is a function of the state at the last planning date $\boldsymbol{s}_{D(i)}$ and of the time since the last planning. Given the Markov assumption, these are sufficient statistics. If the producer chooses a quantity to sell, he or she obtains

$$\Pi^Y(\boldsymbol{s}_{D(i)}, t - D(i)) = \max_{Y_t} E[Q^{-1}(Y_t, \boldsymbol{s}_t) Y_t - C(Y_t, \boldsymbol{s}_t) \mid \Im_t], \tag{2}$$

where $Q^{-1}(Y, \boldsymbol{s}) : \mathbb{R}^{S+1} \to \mathbb{R}$ is the inverse demand function. Since the producer can choose either a price to charge or a quantity to produce, the profits are

$$\Pi(\boldsymbol{s}_{D(i)}, t - D(i)) = \max\{\Pi^P(\boldsymbol{s}_{D(i)}, t - D(i));\ \Pi^Y(\boldsymbol{s}_{D(i)}, t - D(i));\ 0\}.$$

The third possibility allows the firm to shut down if profits are negative.

I make the following assumption on this problem:

Assumption 1. *The functions $C(\cdot, \cdot)$ and $Q(\cdot, \cdot)$ are such that*

i) *The maximization problems leading to $\Pi^P(\boldsymbol{s}, t)$ and $\Pi^Y(\boldsymbol{s}, t)$ are well defined for all $\boldsymbol{s}$ and t; namely, the problems have a solution and expectations can be formed.*
ii) *$\Pi^P(\boldsymbol{s}, t)$ and $\Pi^Y(\boldsymbol{s}, t)$ are finite for all possible $\boldsymbol{s}$.*
iii) *$\Pi(\boldsymbol{s}, t)$ is continuous in both arguments.*

Whenever the agent updates his or her information and plans, he or she incurs a non-negative finite cost given by the continuous function $K(\boldsymbol{s}_t) : \mathbb{R}^S \to \mathbb{R}$. Producers maximize the expected

2. I will denote the expectation conditional on the information at the current planning date by $E[\cdot]$.

present discounted (at the rate $r > 0$) value of profits including planning costs

$$J(s_0, D) = E\left\{\sum_{i=0}^{\infty}\left(\int_{D(i)}^{D(i+1)} e^{-rt}\Pi(s_{D(i)}, t - D(i))dt - e^{-rD(i+1)}K(s_{D(i+1)})\right)\right\} \tag{3}$$

by choosing a sequence of planning dates $D = \{D(i)\}_{i=1}^{\infty}$ that is $\Im$-measurable. Note that if the costs of planning are always 0, the producer optimally chooses to be always attentive.

This problem has a recursive structure between adjustment dates. Letting $\boldsymbol{s}$ denote the state at the current planning date and $\boldsymbol{s}_d$ the state at the next planning date, I can write the problem as

$$V(s) = \sup_d\left\{\int_0^d e^{-rt}\Pi(s, t)dt + e^{-rd}E[-K(s_d) + V(s_d)],\ \text{subject to } s_d = \Psi(s, u^d)\right\}. \tag{4}$$

Because I passed the expectations operator through d, I have imposed the constraint that the date of the next plan must be conditional on the information at the current planning date. Note that one can see the producer as choosing his or her next planning date either at the current planning date or instead at that future date. Since the producer receives no new information while inattentive, his or her choice will be the same regardless of when it is made. Bellman's principle of optimality then implies that:[3]

Proposition 1. *The dynamic program in* (4) *has the same solution as maximizing* (3)*:*

$$V(s) = \sup_D J(s, D).$$

There is a well-defined, continuous, finite, and unique value function solving this problem, and a set of necessary first-order conditions characterizing the solution.

The problem in (4) may strike some readers as similar to optimal stopping problems and their associated regulated Brownian motions. However, in those problems, the producer observes the state of the economy every instant and decides whether to adjust. Adjustment is then *state-contingent*. In the inattentiveness model instead, in between adjustments, the producer is getting no new information. Whereas regulated Brownian motion problems lead to adjustments contingent on the current state of the economy, inattentive agents adjust at optimally chosen dates regardless of the state of the economy at those dates. The optimal planning intervals are not necessarily always the same though, since they depend recursively on the state of the economy at the last adjustment date. Adjustment with inattentiveness is therefore *recursively time-contingent*, independent of the current state, but a function of the state at the last adjustment.

This difference between state-contingent and recursively time-contingent adjustment leads to very different dynamics and predictions. For instance, in the inattentiveness model, today's news do not affect the fraction of producers adjusting today. In state-contingent models instead, a large shock to, for instance, monetary policy leads many producers to adjust immediately, which offsets the real effects of such a shock. Another example of the difference between the two models is that in the inattentiveness model, those who adjust may or may not have been charging a price that was far away from the optimum. The inattentive producers expect to be far away from target, but they may or may not be, depending on the actual current and past news. In state-contingent models, those who adjust are those whose current prices *are* far away from the optimum. As Danziger (1999) and more recently Golosov and Lucas (2003) emphasize, this feature of state-contingent models again attenuates the real impact of monetary policy shocks.

3. The Appendix contains the proof of this and all the other propositions.

3. WHAT TO PLAN

The producer must first choose whether to set a plan for prices or a plan for quantities.[4] An immediate result is the following:

Proposition 2. *If demand is certain, the producer is indifferent between price and quantity plans.*

The proof is straightforward: if the demand function is fixed, then setting a price fixes a quantity, and setting a quantity fixes a price. The producer can choose a price–quantity pair in the stable demand function. Being inattentive may be costly, but it is equally so for price and quantity plans.

Shocks to demand break this equivalence between price and quantity plans, since setting one leaves the other to vary with the shocks to ensure market clearing. With both types of shocks and denoting by Q_x the partial derivative of the demand function with respect to x evaluated at $\mathbf{s} = E[\mathbf{s}]$:

Proposition 3. *Up to a second-order approximation in the size of the shocks $\|\mathbf{s}\|$, producers prefer plans for prices if and only if*

$$Q_s Q_{ps} + \left(-\frac{Q_s^2}{2Q_p}\right) Q_{pp} + \frac{Q_p^2}{2Q}(C_{qq} Q_s^2 + 2C_{qs} Q_s) \leq 0. \tag{5}$$

Producers prefer quantity plans otherwise, and are indifferent in case of equality.

To understand the intuition behind this result, consider the case of a monopolist with a zero marginal cost of production facing a linear demand curve with slope -1 subject to a scalar multiplicative shock with an expected value of 1. The condition for price plans to be preferred becomes $Q_s Q_{ps} < 0$. Graphically, in (Y, P) space, this implies that when it shifts out, the demand curve becomes flatter; when it shifts in, the demand curve becomes steeper. This is depicted in Figure 1. The optimal price is P^* and the optimal quantity is Y^* where the 45° line intersects the demand curve. If a shock shifts demand out, with a price set at P^*, the producer will now sell Y', which raises profits by the area of the rectangle $ABY'Y^*$. With a quantity plan, the producer will sell at price P', and profits increase by the area of $ACP'P^*$. Clearly, price plans raise profits by more if $\overline{AB} > \overline{AC}$. But, since under condition (5), this positive demand shock ($Q_s > 0$) makes the demand curve flatter ($Q_{ps} < 0$), it must be that $\overline{AB} > \overline{AC}$. Conversely, a negative demand shock shifts the demand curve inwards and makes it steeper. A price plan sells Y'' units, while a quantity plan charges P''. Since demand is steeper, $\overline{AD} < \overline{AE}$, so price setting leads to smaller losses. Therefore, if (5) holds, price plans lead to larger gains with positive shocks and smaller losses with negative shocks, so the producer prefers them.

Now let the demand function have some curvature ($Q_{pp} \neq 0$). Figure 2 plots the case of an outward shift in demand with zero marginal costs, but now in the case when $Q_{ps} = 0$ so the slope is unchanged, so we can focus on the second term in (5). According to the proposition, the producer prefers price plans if $Q_{pp} < 0$. From the figure, clearly if the demand function is linear then $\overline{AB} = \overline{AC}$, and the producer is indifferent between the two plans. Fixing the horizontal dislocation of the demand curve after the shock, and letting the demand curve now be concave,

4. Weitzman (1974) asked whether a central planner should fix *ex ante* the demand for a product in terms of price or quantity, knowing the firm will respond to shocks. The problem in this paper is the exact opposite. It is the firm that is committing *ex ante* and demand that is moving with shocks.

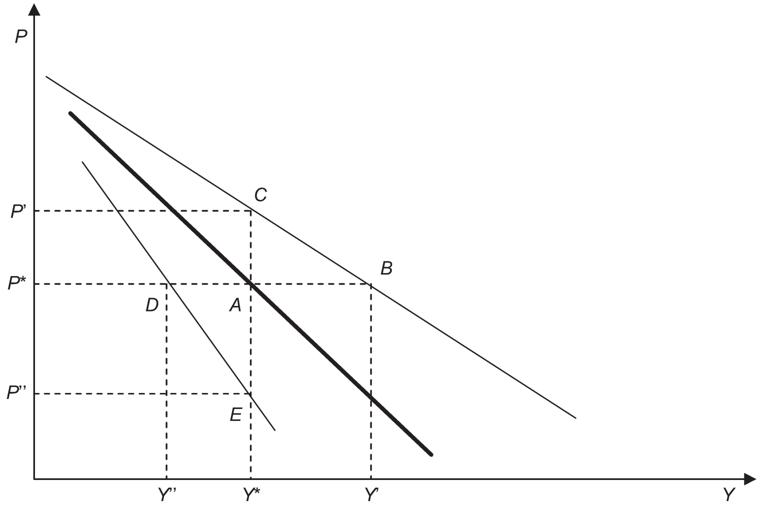

FIGURE 1

Price vs. quantity plans—linear demand

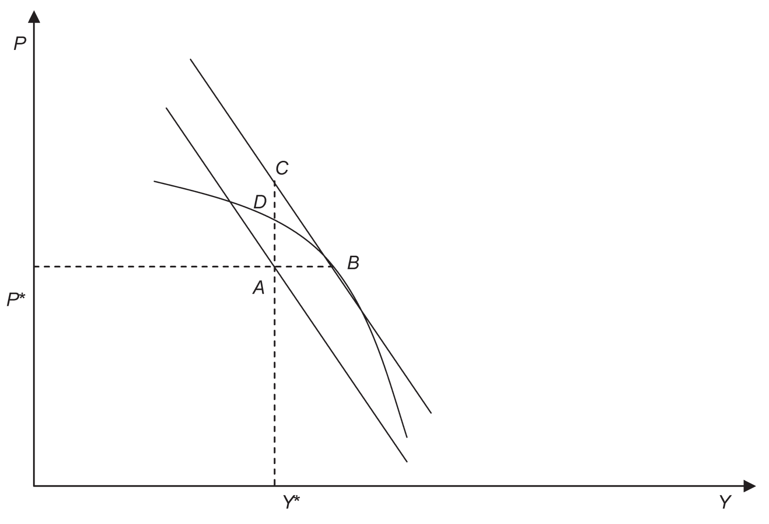

FIGURE 2

Price vs. quantity plans—concave demand

under a quantity plan the price increases only by $\overline{AD}$. Since $\overline{AB} > \overline{AD}$, price plans are preferred. The case of negative shocks works likewise.

Table 1 evaluates Proposition 3 in the case of constant marginal costs c for a few commonly used demand specifications. Notably, with the iso-elastic demand function with multiplicative shocks that is often used in macroeconomics and international economics, price plans are preferred (case (i)). With the logistic specification commonly used in empirical studies of market demand in microeconomics and industrial organization, as long as the constant in the logistic regression is not too large so the firm does not capture a very large amount of the market share,

TABLE 1

Price or quantity plans for different specifications of demand

Case	Demand function	Parameter restrictions	Preferred plan
(i)	$Y = sP^{-\theta}$	$\theta > 1$	Price
(ii)	$Y = \frac{1}{1+e^{bP-s}}$	$b > 0$	Price if $E[s] \leq 2 + bc$ quantity otherwise
(iii)	$Y = f(P) + s$	$f_p < 0$	Price if $f_{pp} \leq 0$ quantity otherwise
(iv)	$Y = sf(P)$	$s > 0, f > 0, f_p < 0$	Price

price plans are also preferred (case (ii)). These cases are fortunate, since casual observation seems to point towards price plans in the world. At least for the common specifications of demand used by economists, the model predicts this should be the case. More generally, if the demand function is subject to either additive or multiplicative shocks, a sufficient condition for price plans to be preferred is that the demand function is concave with respect to price (cases (iii) and (iv)).

The third term in equation (5) involves the slope of marginal costs. The term $C_{qq}Q_s^2$ shows that that decreasing marginal costs ($C_{qq} < 0$) provides an extra incentive for price plans. Intuitively, recall that the optimal quantity sold with full information is determined by marginal costs equalling marginal revenue. If marginal costs are steeply increasing, then shifts in marginal revenue have a small impact on the optimal quantity sold, so that a quantity plan is close to optimal. If instead marginal costs are decreasing, shifts in demand lead to a large discrepancy between the optimal quantity and the one set by a plan and this explains why $C_{qq} < 0$ makes price plans preferred. Klemperer and Meyer (1986) emphasized this effect in their study of whether firms strategically interact as in the Bertrand model or as in the Cournot model. The term $2C_{qs}Q_s$ shows that if an outward shift of demand ($Q_s > 0$) lowers marginal costs ($C_{qs} < 0$) then price plans are preferred. Intuitively, following the shock, a quantity plan leads to a higher price being charged, but since marginal costs are lower, the producer should be charging a lower price. Following a quantity plan is therefore more costly, so a price plan is preferred.

Proposition 3 applies only up to a second-order approximation and, as with all approximations, there is an error that may become large as shocks accumulate. In general, one could use numerical methods to obtain more accurate solutions in some special cases. In a few cases, analytical exact solutions exist, and it is worth solving one of these explicitly. It is common to assume that demand is iso-elastic with multiplicative shocks, $Q(\varepsilon_t, P_t) = \varepsilon_t P_t^{-\theta}$, where ε is a non-negative i.i.d. demand shock and $\theta > 1$ is the elasticity of demand. I further assume that the marginal cost of production, s, follows an independent geometric Brownian motion with volatility $\sigma > 0$ and that planning costs a fixed share κ of profits. The approximate result from Table 1 is that, in this case, price plans are preferred. In this special case, we can solve exactly for expected profits, and find that they are higher with a price plan if $E[\varepsilon]^{1/\theta} \geq E[\varepsilon^{1/\theta}]$. This is true by Jensen's inequality, confirming the approximate result.

Whether in the world we observe price or quantity plans, and whether the choice between them accords with the predictions above, are interesting empirical questions. Using data on which type of plan firms follow, the results in this section and in Klemperer and Meyer (1986) provide a number of testable predictions.[5]

5. The type of plan can either be inferred from time series on prices and output of firms, or it can be collected directly, as in Aiginger (1999), who asked a sample of managers of 930 Austrian manufacturing businesses "What is your main strategic variable: do you decide to produce a specific quantity, thereafter permitting demand to decide upon price conditions, or do you set the price, with competitors and the market determining the quantity sold?" In response, 68% of managers professed to follow price plans, while 32% admitted to quantity plans.

4. THE DETERMINANTS OF INATTENTIVENESS

4.1. *The optimality conditions*

Recall that optimal inattentiveness solves the Bellman equation (4). The necessary first-order condition for optimality is as follows:

$$\Pi(\boldsymbol{s}, d) = E\left[r\left(V(\boldsymbol{s}_d) - K(\boldsymbol{s}_d)\right) + \left(K_s(\boldsymbol{s}_d) - V_s(\boldsymbol{s}_d)\right)\frac{\partial \Psi(\boldsymbol{s}, u^d)}{\partial d}\right]. \tag{6}$$

On the L.H.S. is the flow value from not planning, which equals the profits from keeping to the old plan. On the R.H.S. is the value from planning, which equals the sum of two terms. The first term is the flow value from planning, which is the difference between the value of having a fresh plan and the cost of writing it. The second term is the cost from postponing planning for another instant in which the cost and value of a new plan may change.

The envelope theorem conditions with respect to each component j of the state vector $\boldsymbol{s}$ are

$$V_j(\boldsymbol{s}) = \int_0^d e^{-rt}\Pi_j(\boldsymbol{s}, t)dt + e^{-rd}E[(-K_s(\boldsymbol{s}_d) + V_s(\boldsymbol{s}_d))\Psi_j(\boldsymbol{s}, u^d)]. \tag{7}$$

Equations (4), (6), and (7) characterize the value function $V(\boldsymbol{s})$ and optimal inattentiveness $d(\boldsymbol{s})$.

4.2. *A general approximate solution*

The dynamic program in (4) can be easily solved numerically. Analytically, in general, the optimal inattentiveness is a complicated function of the state of the economy. However, a simple approximate solution can be found by perturbing the problem around the point where the costs of planning are 0. This approach requires only that $V(\boldsymbol{s})$ and $d(\boldsymbol{s})$ are locally differentiable with respect to the costs of planning. Define the function $G(\boldsymbol{s}, t) : \mathbb{R}^{S+1} \to \mathbb{R}$ as the expected difference between profits earned with full information and profits earned while following a pre-chosen plan. Then

Proposition 4. *A perturbation approximation of the optimal inattentiveness around the situation when planning is costless is*

$$d^*(\boldsymbol{s}) = \sqrt{\frac{2K(\boldsymbol{s})}{G_t(\boldsymbol{s}, 0)}}.$$

This solution shows that inattentiveness is determined by two factors. First, the larger the costs of planning are, the longer is inattentiveness. Moreover, since $d^*(\cdot)$ is of order $\sqrt{K}$, second-order costs of planning lead to first-order long inattentiveness. The reason is that inattentive agents are near-rational in the Akerlof and Yellen (1985) sense. While optimal inattentive behaviour differs from optimal behaviour with full information, because the profit function is flat at a maximum, this deviation only has a second-order effect on profits (Mankiw, 1985). The agent is therefore willing to tolerate a first-order period of inattentiveness with only second-order costs of planning, since the inattentiveness involves a loss in profits that is also only second order.[6]

6. In contrast, in state-contingent adjustment models of sticky prices, fourth-order adjustment costs induce first-order rigidities (Dixit, 1991). Since, in these models, the producer observes the current state of the world, he or she perceives an option value of not adjusting since the state might change in the future, making an adjustment unnecessary.

The second determinant of inattentiveness is $G_t(\cdot)$. The faster the losses from being inattentive accumulate, the shorter is inattentiveness. This could be the case if demand or production are very volatile so that larger forecast errors of the future are more likely. Another reason for a large G_t is profits that are very elastic with respect to price or quantity, so that small errors due to inattentiveness lead to large foregone profits.

4.3. *The iso-elastic case*

This case was introduced at the end of Section 3. It assumes that demand is iso-elastic with price elasticity $\theta > 1$ and is subject to i.i.d. multiplicative shocks, while marginal costs follow a geometric Brownian motion with variance σ^2, and planning costs a fixed share κ of profits. In this case, the producer sets a plan for prices, charging $P_t = (\theta/(\theta-1))E[s_t]$.[7] Then, the following result holds:

Proposition 5. *With iso-elastic demand, optimal inattentiveness solves the equation:*

$$2re^{-\frac{\theta(\theta-1)\sigma^2}{2}d^*} - \theta(\theta-1)\sigma^2 e^{-rd^*} + [\theta(\theta-1)\sigma^2 - 2r](1-\kappa r) = 0.$$

This solution is independent of the states of demand or production. If $\kappa > 1/r$, it equals infinity. If $\kappa < 1/r$, then d^* is unique and finite, and it increases with κ, decreases with σ^2, and decreases with θ. In the vicinity of $\kappa = 0$, it approximately equals

$$d^* = \sqrt{\frac{4\kappa}{\sigma^2\theta(\theta-1)}}.$$

This result illustrates the determinants of inattentiveness. First, inattentiveness is larger, the larger are the costs of planning, and it is first-order long with second-order planning costs. Second, more volatile shocks lead to more frequent updating since inattentiveness is more costly in a world that is rapidly changing.[8] Third, a smaller price elasticity of demand implies that the optimal price is less responsive to fluctuations in marginal costs. The inattentive price is therefore, on average, closer to the full information price. The loss from being inattentive is therefore smaller and the agent stays inattentive for longer.

There is some evidence in favour of this last prediction. Bils and Klenow (2004) find that variables capturing the flexibility of demand account for much of the variation in the frequency of price adjustment across goods. For instance, most goods sold in supermarkets and grocery stores have very elastic demands since there is intense competition in these goods from multiple stores and brands. These prices are among those that seem to change more often in response to market conditions. In opposition, consider the 10 most infrequently revised prices in the U.S. according to Bils and Klenow (2004). Four of these are fees set by the government, while another three are coin-operated machines and magazines, for which there are clear high physical costs of changing prices. That the prices of these seven goods are adjusted very infrequently is not mysterious. The other three are more interesting: vehicle inspection, legal fees, and safe deposit box rentals. These are all goods for which demand is likely not very sensitive to prices, thus supporting the prediction of the model.

7. A curious property of the iso-elastic demand function is that the optimal price does not depend on the state of demand. If there were no technology shocks, the producer could be inattentive forever.

8. This prediction distinguishes the inattentiveness model from the limited information capacity model in Moscarini (2004), in which higher volatility can lead to more infrequent adjustments. Future empirical work can test these opposite predictions.

Another piece of evidence in favour of this result comes from the European data on individual prices across sectors reported in Dhyne, Álvarez, Bihan, Veronese, Dias, Hoffman, Jonker, Lünneman, Rumler and Vilmunen (2004). They find that the prices of services are the least frequently adjusted, whereas prices in the energy sector are very frequently updated. Two features of the energy retail sector are that the cost of its raw inputs are subject to frequent shocks (high σ^2) and consumers are very sensitive to price changes (high θ). Goods in the services sector, on the other hand, are typically differentiated, so that their price elasticity of demand is low. The theoretical results in this section therefore seem to be consistent with the broad patterns in the data. Systematically testing these predictions empirically is an interesting topic for future research.

4.4. *Real and nominal rigidities*

It is common in macroeconomics to consider a world in which there are many identical firms indexed by j, each a monopolist setting the price of a good facing a state of the economy composed of the price level, P, the level of aggregate demand, Y, and shocks to productivity, A. The profit function then becomes $\pi(p(j)-p, y, a)$, where lower case letters denote the logarithms of the respective capital letters. The natural level of output, y^n, is defined as the output level if the costs of planning are 0 so all the producers are attentive. The Appendix shows that in this case, a producer in an inattentive economy is inattentive for approximately

$$d^* = \frac{2}{\alpha}\sqrt{\frac{K}{-\pi_{pp}\mathrm{Var}\left[y-y^n\right]}}. \tag{8}$$

An important determinant of optimal inattentiveness is α, which equals $-\pi_{py}/\pi_{pp}$. Ball and Romer (1990) named this last parameter the inverse of an index of "real rigidities". The reason for this label is that a first-order log-linear approximation shows that a producer wishes to set its price equal to $p+\alpha(y-y^n)$. The parameter α therefore measures how much the firm wishes to change its price in response to shocks. If α is small, not responding to shocks is close to being optimal, so being inattentive involves a small cost and producers are inattentive for longer, precisely as we see in equation (8). Longer inattentiveness in turn implies that prices take longer to react to shocks, so a low α is the key property of the profit function that ensures substantial nominal rigidities.

5. AGGREGATION

5.1. *Aggregation with identical firms*

In an economy with many inattentive producers, what can one say about the distribution of their decision dates? At first, one might expect that this distribution depends so tightly on the assumptions about the individual producers, that little can be concluded in general. Surprisingly, it turns out that there are some general answers to this question.

Assume that there are many producers in the economy. The sequence of optimally chosen planning dates for each producer $D=\{D(i)\}_{i=1}^{\infty}$ forms a sequence of stochastic increasing events, while the inattentiveness intervals $\{d(i)\}_{i=1}^{\infty}$ are a sequence of non-negative random variables. I assume that the costs of planning are positive almost surely, so the probability that two or more decision dates occur simultaneously for a given producer is 0. I also assume that planning dates are not always integral multiples of some non-negative number, so D is not a lattice. While this case could be considered, I prefer to focus on the more interesting case where inattentiveness varies randomly with changes in the profits of firms and the costs of planning.

The arrival of decision dates then takes the form of a stochastic point process. Its properties are described by a set of probability density functions for how long the inattentiveness period will last, conditional on when the producer last adjusted. I denote these by $f_i(t)$ and assume that:[9]

Assumption 2. *The densities $f_i(t)$ describe random variables that are*

i) *mutually independent;*
ii) *independent across producers;*
iii) *the same for all producers.*

Independence of decision dates is convenient since then I only need to keep track of when the last decision date for each producer was. The assumption that all producers are independent and alike in turn allows me to interpret $f_i(t)$ as the actual fraction of agents that are revising their plan at a given instant in time. I will therefore refer to this as the distribution of inattentiveness. In turn, the parameter ρ denotes the intensity of attention, defined as the long-run mean number of planning dates in a unit of time: $\rho = 1/E[d(i)]$ as $t \to \infty$.

While Assumption 2 preserves great generality for the results that follow, it does restrict the domain of the problem. For instance, (i) implies that no permanent shocks are allowed to the producer's computational ability. This excludes events such as the introduction of a new accounting system in a firm that allows it to process information at a lower cost from then onwards. The assumption that inattentiveness is independent across firms in turn precludes aggregate shocks to information processing ability, such as for instance the introduction of computers or the Internet. Finally, (iii) precludes the study of the case when some firms, due to better organization, management, or economies of scale, may have lower information processing costs. Note that (iii) is not a crucial assumption: I will relax it later in this section. Parts (i) and (ii) of Assumption 2, on the other hand, are important for the results that follow. One cannot get results without making some minimal assumptions and I leave the task of relaxing these for future research.[10]

To focus on an economy that has settled at a steady state after operating for a long time, I introduce the following:

Definition 1. The distribution of inattentiveness across firms is

(i) *stationary*, if for any $t > 0$ and any $x \geq 0$, the probability of x decision dates in the interval $(a, a+t)$ is the same for all $a \geq 0$;
(ii) an *equilibrium*, if it is the limit of the system as $t \longrightarrow \infty$.

I focus on studying the stationary equilibrium distribution of inattentiveness across firms.

Given this set-up and without any further assumptions, the following remarkable result holds:

Proposition 6. *The only stationary equilibrium distribution of inattentiveness is the exponential distribution with parameter ρ.*

9. The assumption is stated in terms of properties of the optimally chosen decision dates to make its restrictions more transparent. But, it could also be stated in terms of properties of the state vector $\boldsymbol{s}_t$, using the results in Section 4 to map these into properties of the decision dates.

10. Note that assumption (ii) does *not* preclude the existence of aggregate shocks to the demand or technology, such as the productivity shocks in Section 4.4 (or the nominal income shocks that will appear in Section 6), as long as these shocks do not affect the moments that determine inattentiveness. What it does preclude are aggregate shocks that when realized affect everyone's decision of how long to be inattentive for.

The process of arrival of decision dates is, therefore, a Poisson process with parameter ρ. That is, if at any point in time, we survey the producers on how long ago they last planned, we will find that the share not having planned for x periods equals $\rho e^{-\rho x}$. Every instant, the share of firms planning is constant and equal to ρ. This result is fortunate. The exponential distribution is easy to manipulate and its memoryless property allows for tractable aggregation dynamics.

5.2. *Heterogeneous firms*

Now, I relax the requirement that producers are identical. I still require parts (i) and (ii) of Assumption 2, and I further assume that the inattentiveness distribution of each producer is stationary. In this case, I introduce two new assumptions: (1) that as $J \to \infty$, $\sum_{j=1}^{J} \rho(j)$ tends to a finite constant ρ; and (2) that after a decision date, the probability of there not being a new decision date by the same producer at some point in the next Δ-length period, should tend to unity equally for all producers as Δ tends to 0. Both conditions are aimed at diminishing the probability that one producer accumulates a large number of decision dates in a short period of time and dominates the cross-sectional distribution. In this case

Proposition 7. *As $J \to \infty$, the distribution of inattentiveness across firms tends to the exponential distribution with parameter ρ.*

The combination of Propositions 6 and 7 provides a strong case for using the Poisson process to model the arrival of decision dates in the aggregate economy. Some intuition for these results can be found in other common physical phenomena. Consider a large telephone exchange, which receives an incoming stream of pooled telephone calls from many different independent individuals, or consider the places where flying bombs from many different sources hit the south of London during World War II. Another example is the arrival of goals at the many different matches that compose the World Cup soccer tournament. The distribution of phone call arrivals, the spatial distribution of bombs, and the distribution of arrival of World Cup goals are all, essentially, analogous phenomena to the arrival of the decision dates of agents in an inattentive economy.

These analogies are particularly interesting because while it is difficult to measure the inattentiveness of economic agents, these three physical phenomena are easily observed. A well-known statistical regularity is that all of these physical phenomena empirically follow a Poisson process. In turn, these observations motivated Khintchine (1960) to prove a theorem that provides a precise mathematical justification for these facts, of which Proposition 7 is an application. Both mathematics and empirics therefore provide a strong case for exponentially distributed inattentiveness.[11]

6. AN APPLICATION: A MODEL OF INFLATION

6.1. *The model*

Assume that there are many identical firms (a continuum) indexed by j. Each produces a differentiated good facing a constant price elasticity demand function: $Y_t(j) = Y_t(P_t(j)/P_t)^{-\theta}$. They all operate a linear production technology $Y_t(j) = A_t L_t(j)$, that uses $L_t(j)$ units of labour to produce $Y_t(j)$ units of output subject to exogenous stochastic labour productivity A_t. They hire labour in the market paying a real wage $W_t(j)/P_t$. The inverse labour supply function is $\omega(\log(L_t(j)), \log(Y_t))$. It increases with the amount of labour supplied, with an elasticity of ψ,

11. In state-contingent adjustment models, such a general and simple result does not seem to be possible.

and increases with aggregate income, with an elasticity of σ, through a standard income effect that makes agents prefer more leisure in good times. Finally, assume that the costs of planning are a constant fraction κ_j of profits at the time of planning. To satisfy the conditions for the aggregation results in the previous section, the costs of planning are stochastic and the expectation of $\sqrt{\kappa_j}$ conditional on past information is i.i.d. over time and across producers and $\sqrt{\kappa}$ denotes its mean.

Finally, to close the model, I postulate an exogenous stochastic process for the log of nominal income $m_t = p_t + y_t$ (small letters denote the log of a variable). This limits the applicability of the model to study monetary or fiscal policy since it leaves aside the link between direct policy instruments and nominal income. Likewise, the assumption of a labour supply function $\omega(\cdot)$ prevents the use of the model to study fluctuations in consumption or real wages. However, what these assumptions buy is an ability to study inflation and its links to productivity and nominal income in a relatively general setting, since the assumptions that I make are consistent with most existing models of inflation. Moreover, describing the link between nominal income and inflation goes a long way towards understanding the monetary transmission mechanism. While these assumptions narrow the applicability of the model, and perhaps give it an unashamed partial equilibrium flavour, they allow the model to very generally answer questions about inflation dynamics in a tractable general equilibrium set-up.[12]

6.2. *The type of plan and length of inattentiveness*

The theoretical results proven so far can be applied to this problem. The first result is that since demand has a constant price elasticity and is subject to multiplicative shocks, firms will set plans for their prices.

The optimal price charged at time t by a producer that last updated at time D is, up to a first-order approximation, $p_t(j) = E_D[p_t + \alpha(y_t - y_t^n)]$, where α is the index of real rigidities that equals $(\sigma + \psi)/(1 + \theta\psi)$. The natural level of output, y_t^n, is the output in the economy if agents are attentive, which up to a first-order approximation moves in parallel with productivity: $y_t^n - E[y_t] = ((1 + \psi)/(\sigma + \psi))(a_t - E[a_t])$.

The second main theoretical result concerned the optimal choice of the length of inattentiveness. The profit function is of the form $\pi(p_t(j) - p_t, y_t, a_t)$ that Section 4.4 studied, so equation (8) provides an approximation to the average length of optimal inattentiveness:

$$d^* = \frac{2}{\alpha}\sqrt{\frac{\kappa}{\theta(\theta-1)(\psi+1)\mathrm{Var}[y_t - y_t^n]}}. \tag{9}$$

To assess the predictions of the model for inattentiveness, I consider different possible parameter values for θ, σ, and ψ. My preferred choices are $\theta = 10$, since it implies a mark-up of about 11% consistent with the estimates in Basu and Fernald (1997); $\sigma = 1$, so that real wages and real output grow at the same rate in the long run; and $\psi = 1/0{\cdot}15$, to match the estimates of the elasticity of labour supply surveyed in Pencavel (1986). Compared with other research on inflation, these choices differ from those of Chari, Kehoe and McGrattan (2000) only in ψ, which they set at 1·25. Rotemberg and Woodford (1997) use aggregate data to estimate $\theta = 7{\cdot}88$, $\sigma = 0{\cdot}16$, and $\psi = 0{\cdot}47$. Finally, Ball and Romer (1990) set $\sigma = 0$ and calibrate $\theta = 7{\cdot}8$ and $\psi = 6{\cdot}7$.

I use log output per hour to measure y_t^n since with the benchmark $\sigma = 1$, up to a constant $y_t^n = a_t = \log(Y_t/L_t)$. I measure y_t by quarterly real GNP and use an Hodrick–Prescott filter to

12. Reis (2004) studies the behaviour of inattentive consumers so, in principle, one could build a model with both inattentive consumers and producers. For now, I leave this for future work.

TABLE 2

Optimal expected length of inattentiveness in quarters

	Parameter combinations (θ, σ, ψ)			
Costs of planning	Baseline (10, 1, 6·7)	Chari *et al.* (10, 1, 1·25)	Rotemberg–Woodford (7·88, 0·16, 0·47)	Ball–Romer (7·8, 0, 6·7)
0·046	10	13	26	12
0·028	8	10	20	9
0·010	5	6	12	6
0·001	2	2	4	2

isolate the cycle in the output gap. The standard deviation of $y_t - y_t^n$ in the U.S. data from 1954 to 2003 is 0·014.

Finally, one must choose a value for the costs of planning as a share of profits. Zbaracki *et al.* (2004) followed a large U.S. manufacturing firm through its decision process, and estimated how much it cost for this company to set a new price catalogue. A conservative use of their estimates that considers only the costs that are internal to the firm is 4·6% of the company's net margin. However, the accounting definition of the net margin may not be the most adequate measure of profits in this model. Using instead the Zbaracki *et al.* (2004) estimates of the costs of planning as a share of total costs leads to an estimate of 2·8%. I also consider lower costs of planning of 1% and 0·1%.

Table 2 shows the predictions from equation (9) for the average length of inattentiveness in quarters. A first result to take away from the table is that very small costs of planning can lead to considerable inattentiveness. Even when it costs only 0·1% of profits to plan, producers only plan about every 6 months. A second conclusion is that for the baseline parameters and the Zbaracki *et al.* (2004) estimates of the costs of planning, we should expect to see firms changing their plans about every 2 years. The model therefore predicts inattentiveness of a plausible order of magnitude.

One can turn these predictions into a test of the model. Carroll (2003) and Mankiw, Reis and Wolfers (2004) use data on inflation expectations to infer the speed at which information disseminates in the economy. Both estimate an average inattentiveness of about 1 year. For the four different parameter combinations in the columns of Table 2, costs of planning of 0·7%, 0·4%, 0·1%, and 0·5% of profits respectively, would generate this amount of inattentiveness. These costs are consistent with Zbaracki *et al.* (2004), once you take into account plausible measurement errors. The model is therefore consistent with the independent observations on inattentiveness and on the costs of planning.

6.3. *The Phillips curve*

Up to a first-order log-linear approximation, the log price level equals the sum of the logs of prices set by different producers. If the index of the firms, j, stands for how long has it been since the producer last updated his or her plan, then

$$p_t = \int_0^\infty p_t(j) dH(j),$$

where $H(j)$ is the distribution of how long it has been since the last adjustment. The third main theoretical result in this paper can now be used. It states that $H(j)$ tends to the exponential distribution with parameter ρ equal to the inverse of the average length of inattentiveness. Therefore

$$p_t = \rho \int_{-\infty}^{t} e^{-\rho(t-j)} E_j[p_t + \alpha(y_t - y_t^n)]dj.$$

Taking time derivatives and rearranging, inflation is given by[13]

$$\dot{p}_t = \alpha\rho(y_t - y_t^n) + \rho \int_{-\infty}^{t} e^{-\rho(t-j)} E_j[\dot{p}_t + \alpha(\dot{y}_t - \dot{y}_t^n)]dj.$$

This is a continuous-time version of the sticky-information Phillips curve of Mankiw and Reis (2002). As they showed, it has three desirable features that match the existing evidence. First, disinflations always cause recessions (although announced disinflations lead to smaller recessions than announced ones). Second, monetary policy shocks have their maximum impact on inflation with a substantial delay. Third, the change in inflation is positively correlated with the level of economic activity.[14]

Mankiw and Reis (2002) reached this Phillips curve by making three assumptions. First, they assumed that producers are inattentive, only sporadically updating their information sets. Second, they assumed that they set plans for prices and third, they assumed that the arrival of decision dates is a Poisson process. This paper, instead, only assumed that there is a cost of acquiring, absorbing, and processing information. It derived inattentiveness as the optimal response to such costs. It showed the conditions under which producers choose to set plans for prices and it found that in a world with many agents, the distribution of inattentiveness converges to that of a Poisson process. The inattentiveness model provides a micro-foundation for the sticky-information model.

Having this micro-foundation has many advantages. The model can be used to understand other features of producer behaviour aside from pricing, such as for instance the price vs. quantity decision. Moreover, the model provides a unified framework to study different types of behaviour by different agents. It can be applied to study the actions of consumers, investors, or other economic agents. This is beneficial not just from the perspective of having a theory that is parsimonious and widely applicable, but also empirically, since the model generates predictions across many dimensions that can be tested in different ways. A further advantage of having a micro-foundation is that it links the two key reduced-form parameters, α and ρ, to preference and technology parameters, which is helpful in assessing the likely values of these parameters. Moreover, at least since Lucas (1976), economists have hoped that these parameters are structural in the sense that they do not vary across different policy regimes, and so can be used to reliably forecast future inflation.

7. THREE EMPIRICAL TESTS OF THE MODEL

This section tests the model of inflation in three ways. First, I examine whether the model is able to match the second moments characterizing the post-war U.S. inflation. While Mankiw and Reis (2002) find that the sticky-information model reproduces a few key qualitative features of the data, the question I ask here is whether it can fit the data quantitatively. Second, I examine the usefulness of the micro-foundations at generating a model that can forecast inflation out of sample, by comparing its performance with that of autoregressive models.

13. I use the standard notation $\dot{x}_t$ to denote the time derivative of a generic variable x_t.
14. The Calvo (1983) sticky-price model, *per contra*, can fit none of these facts.

TABLE 3

Model vs. data in the post-war U.S.

	Model	Data
S.D.(Δp_t)	0·0059	0·0062
Corr($\Delta p_t, \Delta p_{t-1}$)	0·9961	0·8859
Corr($\Delta p_t, \Delta m_t$)	0·3749	0·4263
Corr($\Delta p_t, \Delta m_{t-1}$)	0·4240	0·3972
Corr($\Delta p_t, \Delta m_{t+1}$)	0·3555	0·3780
Corr($\Delta p_t, \Delta a_t$)	−0·2436	−0·2667
Corr($\Delta p_t, \Delta a_{t-1}$)	−0·2395	−0·2501
Corr($\Delta p_t, \Delta a_{t+1}$)	−0·2067	−0·1619

Note: The notation Δx_t denotes the quarterly change in variable x_t.
The model's predictions were obtained by simulating the model feeding in the empirical innovations to nominal income and productivity. In the data column are the sample moments in the period 1960:1–2003:4.

The third test of the model is more demanding. I ask whether the model can *also* explain the behaviour of inflation during the pre-war U.S. Because monetary policy was very different in this period, this amounts to asking whether the model survives the Lucas (1976) critique.[15]

7.1. *Can the model fit the second moments of post-war inflation?*

The two relevant reduced-form parameters of the model are α and ρ. Using my baseline parameters for θ, σ, and ψ, the implied value of α is 0·11.[16] For ρ, I use the estimates of Carroll (2003) and Mankiw, Reis and Wolfers (2004) and set $\rho = 0{\cdot}25$ implying an average inattentiveness of 1 year, which, following the discussion in Section 6.2, is also consistent with the other micro-parameters.

To specify the stochastic processes for a_t and m_t, I use quarterly U.S. data from 1954:1 to 2003:4. Data for the log output per hour in the non-farm business sector suggests that a_t is a random walk, with a standard deviations of shocks of 0·008. Nominal GNP growth is well described by an AR(1) with autoregressive parameter 0·39 and a standard deviation of shocks of 0·009.

Table 3 uses these parameter values to display the model's predictions for different second moments of inflation. It also shows the equivalent moments in the U.S. data. The model fits the data remarkably well. It closely fits the univariate properties of inflation, its variability and its persistence. Moreover, it matches well the correlation of inflation with nominal income and productivity, both contemporaneously and with one-quarter leads and lags. With only one exception, all of the model's predictions do not differ from the empirical moments by more than 0·05.

15. These tests do not exhaust the set of empirical applications of the model. For instance, I do not estimate the model. This would require overcoming some challenges with maximizing the likelihood function since the model implies a recursive but infinite moving average representation for inflation. Future research will hopefully make progress on this problem.

16. The parameter α plays two crucial roles. First, a small α leads to long periods of inattentiveness and so a small ρ (Section 4.4). Second, keeping ρ fixed, a smaller α generates larger real effects of nominal shocks. The reason is that, the smaller is α, the stronger are strategic complementarities in pricing, so the firms that are adjusting wish to set their individual prices close to those set by non-adjusting firms. Through these two roles, a small α leads to limited adjustment of prices and so large real effects of nominal shocks. Woodford (2003*a*, pp. 163–173) discusses the calibration of α at length and, taking into account both micro and aggregate evidence, he concludes that a value between 0·10 and 0·15 is adequate. Using the Chari *et al.* (2000) parameters, $\alpha = 0{\cdot}17$, the Rotemberg and Woodford (1997) estimates, $\alpha = 0{\cdot}13$, and the Ball and Romer (1990) parameters, $\alpha = 0{\cdot}13$.

TABLE 4

Out-of-sample forecasting performance

	Model		
Forecasting period	Inattentiveness	VAR(2)	AR(2)
2000Q1–2003Q4	0·0037	0·0051	0·0049
1996Q1–2003Q4	0·0031	0·0039	0·0037

Note: Each cell has the one-step ahead forecast mean squared error multiplied by 1000.

The exception is that the model predicts slightly more serial correlation for inflation than what we find in the data. However, if the measurement of inflation in the data is polluted with some classical measurement error, we should expect the model to predict too much persistence. If measurement error accounts for the tiny discrepancy between the standard deviation of inflation in the model and the data (0·0003), then the model would predict that observed inflation would have a serial correlation of 0·8869, very close to what we observe.

7.2. *Can the model forecast inflation out of sample?*

To forecast, aside from the structural parameters, the model requires as inputs the expectations of future inflation and the growth rate of the output gap as of different times in the past. I obtain these from a bivariate VAR on these two variables, and they enter the model through the tightly specified term, which weighs different past expectations by the weights of the exponential distribution.[17]

I evaluate the model's performance at forecasting inflation one-quarter ahead relative to two unrestricted reduced-form models. The first is an unrestricted bivariate VAR of order 2 on inflation and the growth rate of the output gap. Unlike the tightly specific inattentiveness model of inflation, this allows past information to be used freely in forecasting future inflation The second model is an AR(2) of inflation. The order of both models minimizes the Bayesian information criteria.

The models are estimated using information from the start of the sample, the first quarter of 1954, until a closing date. They are then asked to predict inflation one-quarter ahead from that closing date until the end of my sample at the last quarter of 2003. Table 4 compares the mean squared forecast error of the inattentiveness model against the two reduced-form models for two different choices of the closing date: 1995Q4 and 1999Q4. In the first case, the models must predict inflation for eight full years out of sample, and in the second for 4 years. The message of Table 4 is that the model clearly outperforms reduced-form models at forecasting inflation. The improvement is 24% over the best reduced-form model in the shorter forecasting period and 17% for the longer one.

7.3. *Can the model fit the facts on pre-war inflation?*

Before World War I, monetary policy was very different from what it is now: there was no Federal Reserve system and the gold standard dictated monetary policy. Correspondingly, inflation was

17. I take this approach instead of using the simple exogenous processes for productivity and nominal income in the previous section. With this alternative, it would be difficult to distinguish which part of the model's good performance is driven by (only approximately true) assumptions on the exogenous processes, and which part is due to the actual model of inattentiveness.

close to serially uncorrelated, in stark contrast with its high persistence in the post-war period. While the previous sections showed that the model can explain post-war inflation, this section asks whether it can also explain the data from this very different period.

The data for the pre-war period comes at an annual frequency from two sources. Kendrick (1961) provides estimates of output per hour in the non-farm sector from 1889 to 1913.[18] Nominal income and its deflator from 1869 to 1913 come from Romer (1989).[19] The stochastic properties of these two series are markedly different relative to the post-war period: both nominal income growth and the level of productivity are approximately serially uncorrelated.

The model requires knowledge of the quarterly processes for m_t and a_t, however, and many alternatives are consistent with these annual moments. I proceed by opting for the most parsimonious quarterly statistical representations that are consistent with the annual data. I choose a random walk for m_t, since it implies that annual nominal income should be an IMA(1,1) process with a moving average coefficient of 0·24, a specification that the data does not statistically reject. For productivity, I use a quarterly white noise since it implies a serially uncorrelated annual process.

As for the structural parameters, while the U.S. today is certainly very different from what it was at the beginning of the 20th century, there is no clear indication that the elasticity of demand for products or the income and wage elasticities of labour supply were much different then from what they are now. I therefore assume that these micro-parameters have not changed and so α is still 0·11. By keeping this parameter fixed, if I err, I will do so against the model by forcing it to fit two distinct periods with the same parameters. The parameter ρ depends not only on these elasticities but also on the variance of the output gap and on the costs of planning. Romer (1989) found that the standard deviation of detrended output was 31% higher in the pre-war than in the post-war. In this more volatile pre-war world, if the costs of planning were unchanged, producers would plan more often, about once every three quarters. However, it is plausible that the great advances in information technology during the 20th century have reduced the costs of planning. Producers would wish to plan less often in the pre-war world, when planning was more costly. To give some weight to this argument, in the baseline calibration of the model, I will consider two possibilities for inattentiveness, four and three quarters.[20]

Table 5 contrasts the predictions from the model with the data for the 1890–1913 period. It is noticeable from the second column how different the sample moments are from the post-war estimates in Table 3. It would be remarkable to have a model that could fit both periods. The third column has the average predictions of the model and the fourth column has 90% confidence intervals. Table 5 shows that the performance of the model is not quite as successful as in the post-war data. The model underestimates the variability of inflation, and also cannot match the contemporaneous correlations between nominal income, productivity, and inflation. Nevertheless, over the other dimensions, the model does a good job. It predicts about the right amount of persistence of inflation in the data and it captures well the dynamic relation between inflation and lagged and lead nominal income and lagged and lead productivity. The model can therefore match five of the eight moments in the table.

18. In the beginning of the 20th century, agriculture had a large weight in the U.S. economy with man-hours in the farm sector accounting for about 30% of total man-hours. Nevertheless, measuring a_t as output per hour in the non-farm sector or per hour in the whole economy is not important for my purposes: the correlation coefficient between the two series is 0·98.

19. Balke and Gordon (1989) present a different set of estimates of nominal gross national product and its deflator. While the two estimates of nominal income are quite similar, those of inflation have substantial differences. A previous version of this paper contrasted the inattentiveness model with both sets of data, and found that its performance was similar.

20. A previous version of this paper also considered the case where the costs of planning would have risen to make inattentiveness 3-year long. Details are available from the author.

TABLE 5

Model vs. data in the pre-war U.S.

	Data	Model	90% Confidence interval	Measurement error adjusted	Model with three-quarter plans
S.D.(Δp_t)	0·0280	0·0085	0·0063; 0·0111	0·0280	0·0113
Corr$(\Delta p_t, \Delta p_{t-1})$	0·1615	0·2242	−0·1985; 0·5904	0·0208	0·0527
Corr$(\Delta p_t, \Delta m_t)$	0·7508	0·1184	−0·2020; 0·4154	0·0287	0·1742
Corr$(\Delta p_t, \Delta m_{t-1})$	0·2482	0·4044	0·0840; 0·6601	0·0980	0·4911
Corr$(\Delta p_t, \Delta m_{t+1})$	0·0187	−0·0758	−0·4056; 0·2605	−0·0184	−0·0570
Corr$(\Delta p_t, \Delta a_t)$	−0·0916	−0·6510	−0·8349; −0·4347	−0·1577	−0·7189
Corr$(\Delta p_t, \Delta a_{t-1})$	0·2212	0·3233	0·0746; 0·5585	0·0783	0·3547
Corr$(\Delta p_t, \Delta a_{t+1})$	0·2571	0·3234	0·0729; 0·5604	0·0783	0·3557

Notes: The notation Δx_t denotes the annual change in x_t. The data column has sample moments in 1890–1913. The predictions of the model come from drawing 10,000 innovations to nominal income and productivity from normal distributions with variances set to equal the data, and reporting the average and the 5% and 95% percentiles. The second to last column has the model's predictions assuming that classical measurement error accounts for the difference between actual and predicted inflation. The last column has the model's predictions if average inattentiveness is three quarters.

One possible source of bias is the extent of measurement error in the inflation data. The fifth column in Table 5 reports the predictions of the model assuming that there is classical measurement error of inflation accounting for the discrepancy between the model's predicted standard deviation and the data. This modification reconciles the predictions of the model with the empirical contemporaneous correlation between inflation and productivity. Finally, column 6 reports the predictions of the model when agents are inattentive for three quarters on average. The model's performance is similar to the one with four-quarter plans.

The results in Table 5 are therefore mixed. The model fits some dimensions of the pre-war U.S. data, but misses other features of the data. Given the tall order put forward to the model though, the results are encouraging. Few (if any) of the existing models of inflation would perform this well across such different periods in history.

8. CONCLUSION

I have presented a model in which producers face costs of acquiring, absorbing, and processing information. Producers optimally choose to be inattentive to current news, only sporadically updating their information, expectations, and plans. I derived three main theoretical results. First, I established the conditions under which producers set plans for the price to charge, rather than the quantity to sell. Second, I characterized the determinants of the optimally chosen inattentiveness. Third, I showed that in a large population the exponential distribution should approximate well the distribution of inattentiveness.

This set of results should be useful in constructing models of inattentive economies to study different phenomena. In this paper, I applied the model to study inflation. I showed that the inattentiveness model provides a micro-foundation to the sticky-information Phillips curve. I then tested this micro-founded model of inflation dynamics on three aspects of the data and found that the model performed well. First, the model could replicate very closely the second moments in the post-war data. Second, it beat the reduced-form autoregressive model in out-of-sample forecasting, and third, it fared moderately well at fitting the pre-war facts on inflation, a very demanding test of its invariance across policy regimes.

The inattentiveness model follows in the tradition of the menu cost models introduced by Mankiw (1985). The research that followed his work, however, interpreted menu costs as physical fixed costs of changing prices, leading to an emphasis on sticky-price models. The inattentiveness model instead stresses an interpretation of menu costs as fixed costs of acquiring information, and especially of absorbing and processing it. Plans and information are then sticky, rather than prices. This change in interpretation may seem slight, but it turns out to imply a very different model and implications for inflation dynamics.

The inattentiveness model in this paper is certainly not the only economic model to treat information as a costly good, and the treatment in this paper is admittedly coarse. For instance, the model side-steps some interesting questions that arise when producers look to each other to infer information. (Caballero, 1989, takes a first pass at this problem.) Likewise, while the inattentiveness model emphasizes people's limited ability to absorb and process information, many interesting behavioural questions remain on how to model the details of these limitations. These are fascinating research areas for future work to explore.

Still, for now, this paper has provided the foundations to build models of nominal rigidity based on limited information that provide a counterpart to the micro-founded sticky-price model of Sheshinski and Weiss (1977, 1982), Caballero and Engel (1991), and Caplin and Leahy (1997). Moreover, the tools and lessons in these paper, combined with those in Reis (2004), who studies inattentive consumption choices, suggest that enough progress has been made that it is within our grasp to construct fully fledged, micro-founded, general equilibrium models of interacting inattentive agents. This is not an easy task, and there remain several difficult (but interesting) obstacles to overcome. Given the success that models based on inattentiveness have in describing the data, this seems to be a worthy pursuit.

APPENDIX

Proof of Proposition 1. Since $\Pi(\boldsymbol{s},t)$ and $K(\boldsymbol{s})$ are well defined and continuous and D satisfies the measurability restrictions, then $J(\boldsymbol{s},D)$ is well defined. From Assumption 1, $0 \leq \Pi(\boldsymbol{s},t) < +\infty$ for all $\boldsymbol{s}$ and t. The costs of planning are also non-negative and finite. Therefore $J(\boldsymbol{s},D)$ is bounded below and above. The constraint set for D including the measurability restrictions and the law of motion for the state is clearly non-empty. Bellman's principle of optimality (Stokey and Lucas, 1989) then shows that $V(\boldsymbol{s}) = \max_D J(\boldsymbol{s},D)$. Since $J(\boldsymbol{s},D)$ is well defined and bounded above, so is $V(\boldsymbol{s})$. The fact that $V(\boldsymbol{s})$ exists, is unique, and continuous follows from the continuity of $\Pi(\boldsymbol{s},t)$ and $K(\boldsymbol{s})$, and the fact that $V(\boldsymbol{s})$ is the fixed point of a contraction mapping of continuous into continuous functions (Stokey and Lucas, 1989). ||

Proof of Proposition 3. With full information on the state of the demand shocks $\boldsymbol{s}$, let the optimal choices of price and quantity be denoted by the functions $P(\boldsymbol{s})$ and $Y(\boldsymbol{s})$. These are the solutions from maximizing either $\pi^P(P,\boldsymbol{s})$ with respect to P, or $\pi^Y(Y,\boldsymbol{s})$ with respect to Y. With full information, they are, of course, equivalent: $Y(\boldsymbol{s}) = Q(P(\boldsymbol{s}),\boldsymbol{s})$.

With inattentiveness, define the profit functions: $\pi^P(P_t,\boldsymbol{s}_t) \equiv P_t Q(P_t,\boldsymbol{s}_t) - C(Q(P_t,\boldsymbol{s}_t),\boldsymbol{s}_t)$ and $\pi^Y(Y_t,\boldsymbol{s}_t) \equiv Q^{-1}(Y_t,\boldsymbol{s}_t)Y_t - C(Y_t,\boldsymbol{s}_t)$. Then, the optimal price charged in a price plan is P^*, which solves $\max_{P_t} E[\pi^P(P_t,\boldsymbol{s}_t)]$ and so is implicitly defined by the first-order condition $E(\pi^P_p(P^*,\boldsymbol{s})) = 0$. A first-order Taylor approximation of this equation around $E(\boldsymbol{s})$ shows that $P^* = P(E[\boldsymbol{s}]) + O(\|\hat{\boldsymbol{s}}\|^2)$. I denote $\boldsymbol{s} - E[\boldsymbol{s}]$ by $\hat{\boldsymbol{s}}$. This is the well-known certainty equivalence result that, up to a first-order approximation, optimal choices are equal to the choices with full information if the random variables equal their expected values. By a similar argument, the optimal quantity sold with a quantity plan is $\hat{Y} = Y(E[\boldsymbol{s}]) + O(\|\hat{\boldsymbol{s}}\|^2)$.

Then, note that

$$\begin{aligned}\pi^P(P^*, E(\boldsymbol{s})) &= \pi^P(P(E[\boldsymbol{s}])) + O(\|\hat{\boldsymbol{s}}\|^2, E[\boldsymbol{s}])\\ &= \pi^P(P(E[\boldsymbol{s}]), E[\boldsymbol{s}]) + \pi^P_p(P(E[\boldsymbol{s}]), E[\boldsymbol{s}])O(\|\hat{\boldsymbol{s}}\|^2) + O(\|\hat{\boldsymbol{s}}\|^3)\\ &= \pi^P(P(E[\boldsymbol{s}]), E[\boldsymbol{s}]) + O(\|\hat{\boldsymbol{s}}\|^3),\end{aligned}$$

showing that when $\boldsymbol{s}$ equals its expected value, profits under a price plan differ from profits with full information by at most a third-order term. The second line follows from a Taylor approximation, and the third line from the first-order condition. Similar steps show that $\pi^Y(\hat{Y}, E[\boldsymbol{s}]) = \pi^Y(Y(E[\boldsymbol{s}]), E[\boldsymbol{s}]) + O(\|\hat{\boldsymbol{s}}\|^3)$, so $\pi^P(P^*, E[\boldsymbol{s}]) - \pi^Y(Y(E[\boldsymbol{s}]), E[\boldsymbol{s}])$ is at most third order.

A second-order approximation of the difference between profits with price or quantity plans around $E[\boldsymbol{s}]$ gives

$$\pi^P(P^*,\boldsymbol{s})-\pi^Y(\hat{Y},\boldsymbol{s})=\pi^P-\pi^Y+(\pi_s^P-\pi_s^Y)(\boldsymbol{s}-E[\boldsymbol{s}])+\frac{1}{2}(\pi_{ss}^P-\pi_{ss}^Y)(\boldsymbol{s}-E[\boldsymbol{s}])^2+O(\|\hat{\boldsymbol{s}}\|^3). \tag{10}$$

All the functions on the R.H.S. are evaluated at $(P^*, E[\boldsymbol{s}])$ or $(\hat{Y}, E[\boldsymbol{s}])$. Consider each of the terms in turn. The previous paragraph showed that $\pi^P-\pi^Y$ is of order $O(\|\hat{\boldsymbol{s}}\|^3)$. The second term disappears after taking expectations. As for the third term, using the definitions of the profit functions, $\pi_{ss}^P=P^*Q_{ss}-C_{qq}Q_s^2-C_qQ_{ss}-2C_{qs}Q_s-C_{ss}$ and $\pi_{ss}^Y=Q_{ss}^{-1}Q-C_{ss}$, so since $P^*=P(E(\boldsymbol{s}))+O(\|\hat{\boldsymbol{s}}\|^2)$, it becomes

$$\frac{1}{2}(PQ_{ss}-Q_{ss}^{-1}Q-C_{qq}Q_s^2-C_qQ_{ss}-2C_{qs}Q_s)(\boldsymbol{s}-E[\boldsymbol{s}])^2.$$

Finally, use the first-order condition for profit maximization with respect to prices, $Q+PQ_p=C_qQ_p$, to replace for P. Since price plans are preferred to quantity plans if $\Pi^P(\boldsymbol{s},t)\geq\Pi^Y(\boldsymbol{s},t)$, taking expectations of (10), this condition becomes

$$-\frac{Q}{2Q_p}(Q_{ss}+Q_pQ_{ss}^{-1})-\frac{1}{2}(C_{qq}Q_s^2+2C_{qs}Q_s)\geq O(\|\hat{\boldsymbol{s}}\|^3).$$

Using the inverse function theorem, it is easy to show that $Q_{ss}+Q_pQ_{ss}^{-1}=\frac{2Q_sQ_{ps}}{Q_p}-\frac{Q_{pp}Q_s^2}{Q_p^2}$ so that price plans are preferred if

$$-\frac{Q}{Q_p^2}\left(Q_sQ_{ps}-\frac{Q_{pp}Q_s^2}{2Q_p}\right)-\frac{1}{2}(C_{qq}Q_s^2+2C_{qs}Q_s)\geq O(\|\hat{\boldsymbol{s}}\|^3).$$

Rearranging gives the condition in Proposition 3. ‖

Proof of Proposition 4. Rewrite the costs of planning as $K(\boldsymbol{s}_t)=\kappa^2\tilde{K}(\boldsymbol{s}_t)$, where κ is a non-negative scalar. I will approximate the solution around $\kappa=0$.[21] First, subtract the discounted profits obtained from setting prices or quantities with current information on $\boldsymbol{s}_t$. Using $V(\cdot)$ to denote the value function for this problem (a slight abuse of notation):

$$V(\boldsymbol{s})=\max_d\left\{-\int_0^d e^{-rt}G(\boldsymbol{s},t)dt+e^{-rd}E[-\kappa^2\tilde{K}(\Psi(\boldsymbol{s},u^d))+V(\Psi(\boldsymbol{s},u^d))]\right\}. \tag{11}$$

The optimality conditions are only slightly different:

$$-G(\boldsymbol{s},d)+r\kappa^2E[\tilde{K}(\boldsymbol{s}_d)]=E\left[rV(\boldsymbol{s}_d)+(\kappa^2\tilde{K}_s(\boldsymbol{s}_d)-V_s(\boldsymbol{s}_d))\frac{\partial\Psi(\boldsymbol{s},u^d)}{\partial d}\right], \tag{12}$$

$$V_j(\boldsymbol{s})=-\int_0^d e^{-rt}G_j(\boldsymbol{s},t)dt+e^{-rd}E[(-\kappa^2\tilde{K}_s(\boldsymbol{s}_d)+V_s(\boldsymbol{s}_d))\Psi_j(\boldsymbol{s},u^d)], \tag{13}$$

$$V_\kappa(\boldsymbol{s})=e^{-rd}E[-2\kappa\tilde{K}(\boldsymbol{s}_d)+V_\kappa(\boldsymbol{s}_d)]. \tag{14}$$

The last condition is the envelope theorem condition with respect to κ.

The system of equations (11)–(14) defines the optimum. When $\kappa=0$, the solution to the system is $d^*=0$ and $V(\boldsymbol{s})=0$. At this optimum, $G(\boldsymbol{s},0)=0$ for all $\boldsymbol{s}$ and $G_j(\boldsymbol{s},0)=0$ as well. Similarly, the n-th-order derivatives of V with respect to $\boldsymbol{s}$ are all 0. Perturbing the system (11)–(14) by differentiating with respect to κ and evaluating at $\kappa=0$ (where $d^*=0$, $V=0$, $V_s=0$):

$$\begin{aligned}
V_\kappa &= V_\kappa\\
-G_td_\kappa &= rV_\kappa-\frac{d}{d\kappa}\left[\frac{1}{dt}E(dV)\right]\\
V_{j\kappa} &= V_{s\kappa}\Psi_j\\
0 &= -2\tilde{K}-rV_\kappa d_\kappa+\frac{d}{d\kappa}\left[\frac{1}{dt}E(dV)\right]d_\kappa.
\end{aligned}$$

All the functions are evaluated at $\boldsymbol{s}$ and $t=0$. The first and third equations contain no information but the second and fourth form a system of equations that I can use to, substituting for the term in $E(dV)$, solve for d_κ:

$$d_\kappa=\sqrt{\frac{2\tilde{K}}{G_t}}.$$

21. The reader is invited to check that perturbing with respect to κ^2 leads to a bifurcation. The method of undetermined gauges could be used to show that κ is the leading term in the approximation.

Since the approximation to d^* is $d^* = d_\kappa \kappa$, and since $\sqrt{K} = \kappa\sqrt{\tilde{K}}$, the expression for d^* follows. ||

Proof of Proposition 5. The optimal price can be found by maximizing expected profits. Using this optimal price to evaluate the profit function shows that expected profits under a price plan are

$$\Pi(s,t) = \underbrace{\frac{E[\varepsilon]}{\theta-1}\left(\frac{\theta}{\theta-1}\right)^{-\theta}}_{\equiv\Xi} E[s_t]^{1-\theta} = \Xi s^{1-\theta},$$

where the second equality follows from the fact that $E(s_d) = s$ for a geometric Brownian motion. The problem in (4) is in this case

$$V(s) = \max_d \left\{ \Xi s^{1-\theta} \int_0^d e^{-rt} dt + e^{-rd} E[-\kappa \Xi s_d^{1-\theta} + V(s_d)] \right\}.$$

Given the iso-elastic form of the return function, the value function is iso-elastic as well. Let $V = As^{1-\theta}$, where A is a coefficient to be determined. The Bellman equation then becomes

$$As^{1-\theta} = \max_d \left\{ \frac{\Xi s^{1-\theta}(1-e^{-rd})}{r} + e^{-rd}(-\kappa\Xi + A)E[s_d^{1-\theta}] \right\}.$$

Cancelling terms and since $E(s_d^{1-\theta}) = s^{1-\theta} e^{bd}$, where $b = 0.5\theta(\theta-1)\sigma^2$ as s_t is a geometric Brownian motion

$$A = \max_d \left\{ \frac{\Xi(1-e^{-rd})}{r} + e^{(b-r)d}(-\kappa\Xi + A) \right\}. \tag{15}$$

The first-order condition from the maximization problem is

$$\frac{\partial A}{\partial d} = e^{-rd}[\Xi + (b-r)e^{bd}(-\kappa\Xi + A)] = 0.$$

At the optimum d^*, (15) gives the solution for A:

$$A = \frac{\Xi(1-e^{-rd^*}) - r\kappa\Xi e^{(b-r)d^*}}{r(1-e^{(b-r)d^*})}.$$

Using this in the first-order condition and rearranging then yields the condition

$$\Gamma(b,\kappa,d^*) \equiv re^{-bd^*} - be^{-rd^*} + (b-r)(1-\kappa r) = 0.$$

Substituting for b and multiplying by 2 gives the result in the proposition.

Next, I check the second-order conditions for the maximization problem in (15). Note that

$$\frac{\partial^2 A}{\partial d^2} = -r\underbrace{e^{-rd}[\Xi + (b-r)e^{bd}(-\kappa\Xi + A)]}_{=\partial A/\partial d} + e^{-rd} b(b-r)e^{bd}(-\kappa\Xi + A).$$

At the optimal d^*, the first-order condition implies that the first term in the sum is 0 and that the second term equals $-\Xi be^{-rd}$. Therefore,

$$\frac{\partial^2 A}{\partial d^2} = -\Xi be^{-rd} < 0,$$

which guarantees that the zero of the function $\Gamma(b,\kappa,d)$ corresponds to a maximum.

The optimal choice of inattentiveness d^* is the zero of $\Gamma(\cdot)$. Consider then two cases: (i) $b > r$, and (ii) $r > b$. In case (i), it is easy to show that for $\kappa > 0$, then $\Gamma(b,\kappa,0) < 0$, $\Gamma_d(\cdot) > 0$, and $\lim_{d\to\infty}\Gamma(\cdot) = (b-r)(1-\kappa r)$. It follows that if $\kappa < 1/r$ there is a unique optimal finite d; otherwise $d^* = +\infty$. For $d^* > 0$, the implicit function theorem implies that $\text{sign}\{\partial d^*/\partial\kappa\} = \text{sign}\{-\Gamma_\kappa(\cdot)\}$, which is positive so that d^* increases with κ. Similarly, it takes a little work to show that $\Gamma_b(b,\kappa,d^*) > 0$, which implies that d^* decreases with b, and therefore with σ^2 and θ. Turn now to case (ii). Now $\Gamma(b,\kappa,0) > 0$ and $\Gamma_d(\cdot) < 0$ but still, if $\kappa < 1/r$, then $\lim_{d\to\infty}\Gamma(\cdot) < 0$, so there is a unique finite d^*. It is easy to show that now $\Gamma_\kappa(\cdot) > 0$ while it takes some work to show that $\Gamma_b(b,\kappa,d^*) < 0$, from where the same comparative statics follow.

Finally, to obtain the approximation, you can use the result in Proposition 4, but the condition $\Gamma(b,\kappa,d^*)=0$ allows for a check on this result. Let $\tilde{\kappa}=\sqrt{\kappa}$, and note that $\Gamma(b,0,0)=0$, that $\Gamma_d(b,0,0)=0$ and that $\Gamma_{\tilde{\kappa}}(b,0,0)=0$. The implicit function theorem, $\Gamma_d d_{\tilde{\kappa}}+\Gamma_k=0$ then does not apply since $\Gamma_d=0$ and $\Gamma_{\tilde{\kappa}}=0$ so the point $\tilde{\kappa}=d=0$ is a bifurcation point. One further round of differentiation plus the fact that $\Gamma_{d\tilde{\kappa}}(b,0,0)=0$ lead to the conclusion that $d_{\tilde{\kappa}}=\sqrt{-\Gamma_{\tilde{\kappa}\tilde{\kappa}}/\Gamma_{dd}}$. A little more algebra shows that $\Gamma_{\tilde{\kappa}\tilde{\kappa}}(b,0,0)=-2r(b-r)$ and $\Gamma_{dd}(b,0,0)=br(b-r)$. Since a first-order Taylor approximation of d^* around $\tilde{\kappa}=0$ is given by $d^*=d_{\tilde{\kappa}}\sqrt{\kappa}$, the result in the proposition follows. ||

Proof of equation (8). If the agent is inattentive, then he or she will set the same price that all other inattentive agents set. Then $p(j)=p$, which solves $E[\pi_p(0,y,a)]=0$. If the agent is attentive, then he or she sets price $p(j)^*$ that solves: $\pi_p(p(j)^*-p,y,a)=0$. A second-order approximation around the point $(0, E[y], E[a])$ of the difference between profits if attentive or inattentive is

$$\begin{aligned}
&\pi(p(j)^*-p,y,a)-\pi(0,y,a)\\
&\quad=\pi+\pi_p(p(j)^*-p)+\pi_y(y-E[y])+\pi_a(a-E[a])+\frac{1}{2}[\pi_{pp}(p(j)^*-p)^2+\pi_{yy}(y-E[y])^2+\pi_{aa}(a-E[a])^2]\\
&\qquad+\pi_{py}(p(j)^*-p)(y-E[y])+\pi_{pa}(p(j)^*-p)(a-E[a])+\pi_{ya}(y-E[y])(a-E[a])\\
&\qquad-\pi-\pi_y(y-E[y])-\pi_a(a-E[a])-\frac{1}{2}[\pi_{yy}(y-E[y])^2+\pi_{aa}(a-E[a])^2]-\pi_{ya}(y-E[y])(a-E[a]).
\end{aligned}$$

All the functions are evaluated at $(0, E[y], E[a])$. Cancelling common terms and since $\pi_p=0$,

$$\pi(p(j)^*-p,y,a)-\pi(0,y,a)=\frac{1}{2}[\pi_{pp}(p(j)^*-p)^2+2\pi_{py}(p(j)^*-p)(y-E[y])+2\pi_{pa}(p(j)^*-p)(a-E[a])].$$

The natural level of output is defined by $\pi_p(0,y^n,a)=0$: it is the output that prevails if all are attentive. A log-linear approximation shows that $\pi_{py}(y^n-E[y])=-\pi_{pa}(a-E[a])$. A log-linear approximation to the first-order condition for $p(j)^*$ gives $p(j)^*-p=\alpha(y-y^n)$, where $\alpha=-\pi_{py}/\pi_{pp}$.

Using these results to substitute for $(a-E[a])$ and for $p(j)^*-p$ in the expression above gives

$$\pi(p(j)^*-p,y,a)-\pi(0,y,a)=-\frac{\pi_{pp}\alpha^2(y-y^n)^2}{2}.$$

From the definition of the $G(s,t)$ function, it then follows

$$G(\boldsymbol{s},t)=-\frac{\pi_{pp}\alpha^2E[(y_t-y_t^n)^2]}{2}.$$

Since $\mathrm{Var}[y-y^n]=E[(y-y^n)^2]-(E[y]-E[y^n])^2$ and since the equation defining the natural level of output implies that $E[y]-E[y^n]$ is second order, it follows that $E[(y_t-y_t^n)^2]=\mathrm{Var}[y_t-y_t^n]$. Finally, $G_t(\boldsymbol{s},0)$ is the instantaneous variance of the output gap. ||

Aggregation: The arrival of decision dates is a point process of the type that is studied in renewal theory. Cox (1962) and Khintchine (1960) are classic references, while Ross (1983) has a more recent treatment. The proofs that follow combine results from this literature. Throughout, I use the following notations:

$\Phi_i(t)$—the cumulative density function associated with $f_i(t)$,
$H_i(t)$—the probability that have not planned since date $D(i-1)$, that is, $H_i(t)=1-\Phi_i(t)$,
$A_i(t)$—the age of a plan, that is, $A_i(t)=t-D(i-1)$ for $D(i-1)\le t<D(i)$,
$V_i(t)$—the remaining duration of the plan, that is, $V_i(t)=D(i)-t$ for $D(i-1)\le t<D(i)$,
$I_i(t)$—the number of plans made by date t, that is, $I(t)=\{i: D(i)\le t<D(i+1)\}$,
$M(t)$—the mean number of planning dates until t, that is, $M(t)=E[I(t)]$.

Proof of Proposition 6. This proof proceeds over a sequence of steps.

Step 1. *Reducing the problem to only two distributions.*

The first step is to reduce the problem of characterizing the infinite set of distributions $\Phi_i(t)$, to that of characterizing only two distributions. The intuition behind this result is that if the economy has converged to a stationary equilibrium distribution and one observes it at a particular instant in time, there are only two distributions characterizing inattentiveness from then on: one for the length until the immediate next planning date, and the other being the stationary distribution. The proof proceeds as follows: define the probability $h(\tau,t)$ for two consecutive periods of length τ and t, respectively, as the probability that (a) there was at least one decision date in τ, (2) there were 0 decision dates in period t. The probability of (b) conditional on (a) is $h(\tau,t)/\Phi(\tau)$. As $\tau\to0$, this is Palm's function $\varphi(t)=\lim_{\tau\to0}h(\tau,t)/\Phi(\tau)$, which gives the conditional probability that no decision dates occur in period t, if the first instant of this period was a decision date. Khintchine (1960, pp. 45–48) proves the following result:

Theorem 1. $\Phi_i(t) = 1 - \varphi(t)$ *for all* $i \geq 2$.

I therefore only need to describe two distributions, $\Phi_1(t)$, and $\Phi(t) = \Phi_i(t)$ for all $i \geq 2$.

Step 2. *Proving the Elementary Renewal Theorem:* $\rho = \lim_{t\to\infty} M(t)/t$.

This is one of the most fundamental results in renewal theory. One of many possible versions of a proof follows. From the definition of $D(i)$ and $I(t)$, it follows that

$$\sum_{i=1}^{I(t)+1} d(i) = D(I(t)+1) > t.$$

Since the $d(i)$ are independent, Wald's theorem implies $E\left[\sum_{i=1}^{I+1} d(i)\right] = E[I+1]E[d(i)]$. Taking expectations of the expression above, and using the definition of $M(t)$, gives the condition:

$$\frac{M(t)+1}{t} > \frac{1}{E[d(i)]}.$$

Taking the limit

$$\lim_{t\to\infty}\inf \frac{M(t)}{t} \geq \frac{1}{E[d(i)]}.$$

Next, fix a constant X and define an alternative decision process by

$$\bar{d}(i) = \begin{cases} d(i), & \text{if } d(i) \leq X \\ X, & \text{if } d(i) > X \end{cases}$$

for $i = 1, 2, \ldots$. This in turn defines $\bar{D}(i) = \sum_{j=1}^{i} \bar{d}(j)$, $\bar{I}(t) = \sup\{i : \bar{D}(i) \leq t\}$ and $\bar{M}(t) = E[\bar{I}(t)]$. Since the inattentiveness lengths are bounded above by X

$$\bar{D}(I(t)+1) < t + X.$$

After taking expectations, using Wald's theorem, and taking the limit

$$\lim_{t\to\infty}\sup \frac{\bar{M}(t)}{t} \leq \frac{1}{E[\bar{d}(i)]}.$$

Finally, note that $\bar{D}(i) \leq D(i)$ necessarily, and so $\bar{I}(t) \geq I(t)$. It then must be that $\bar{M}(t) \geq M(t)$. Letting $X \to \infty$, so that $E[\bar{d}(i)] \to E[d(i)]$, we obtain

$$\lim_{t\to\infty}\sup \frac{M(t)}{t} \leq \frac{1}{E[d(i)]}.$$

I have then shown that

$$\lim_{t\to\infty} \frac{M(t)}{t} = \frac{1}{E[d(i)]} = \rho.$$

This proof assumed that $E[d(i)] < \infty$. Otherwise, a similar proof holds using the truncated process above. Since $E[\bar{d}(i)] \to \infty$, then $\rho \to 0$.

Step 3. *Finding the distribution* $f_1(t)$.

I am now ready to find the first of the two distributions. From the definition of V_t, the time until the next planning date is:

$$\text{Prob}(V_\tau = t) = f_1(\tau + t) + \int_0^\tau f(t+u)dM(\tau - u).$$

This is because for the time to the next planning date to be in $(t, t+\Delta t)$, either the first decision date took place in this interval, or the last decision date occurred at some other date u. Take the limit of this expression as $\tau \to \infty$, having the first term go to 0 (which I will verify later). Then, by the elementary renewal theorem

$$\lim_{\tau\to\infty} \text{Prob}(V_\tau = t) = \rho \int_0^\infty f(t+u)du = \rho(1 - \Phi(t)).$$

The second equality uses the definition of $\Phi(t)$. It is important to note that

$$\lim_{\tau\to\infty}\int_0^{\tau} f(t+u)dM(\tau-u)=\rho\int_0^{\infty} f(t+u)du$$

in many cases under only the elementary renewal theorem, but in general it may require a closely related alternative called the key renewal theorem (Ross, 1983, pp. 61–65).

Then, recall that since I am focusing on an equilibrium, time 0 corresponds to an observation of a world that has been operating since $-\infty$. Therefore,

$$f_1(t)=\lim_{\tau\to\infty} Prob(V_\tau=t)=\rho(1-\Phi(t)).$$

Step 4. *Proving that $M(t)=\rho t$.*

Combining steps 2 and 3, one can then characterize the mean number of planning dates in equilibrium. This is the key step in the proof of the proposition. Using the definition of $M(t)$

$$M(t)=\sum_{i=1}^{\infty} i\text{Prob}[I(t)=i]=\sum_{i=1}^{\infty} i(\text{Prob}[D(i)\le t]-\text{Prob}[D(i+1)\le t]).$$

But, $\text{Prob}[D(i)\le t]=\Phi_1*\Phi_{i-1}(t)$ where $*$ stands for a convolution. Then

$$\begin{aligned}
M(t)&=\sum_{i=1}^{\infty} i(\Phi_1*\Phi_{i-1}(t)-\Phi_1*\Phi_i(t))\\
&=\Phi_1+\sum_{i=1}^{\infty}(i+1)\Phi_1*\Phi_i(t)-\sum_{i=1}^{\infty} i\Phi_1*\Phi_i(t)\\
&=\sum_{i=1}^{\infty}\Phi_1*\Phi_{i-1}(t).
\end{aligned}$$

The Laplace transform of it is (using the fact that $\Phi_i(t)=\Phi(t)$ for $i\ge 2$ from step 1)

$$\mathcal{L}(M(s))=\frac{\mathcal{L}(\Phi_1(s))}{1-\mathcal{L}(\Phi(s))}.$$

The Laplace transform of the initial distribution is

$$\mathcal{L}(\Phi_1(s))=\frac{\mathcal{L}(f_1(s))}{s}=\frac{\mathcal{L}(\rho(1-\Phi(s)))}{s}=\frac{\rho(1-\mathcal{L}(\Phi(s)))}{s}$$

where the first and third equalities are standard results for Laplace transforms, and the second equality follows from the result in step 3. Substituting for $\mathcal{L}(\Phi_1(s))$ in the expression for $\mathcal{L}(M(s))$

$$\mathcal{L}(M(s))=\rho/s.$$

Inverting the Laplace transform, it follows that $M(t)=\rho t$.

Step 5. *Proving that the distribution is exponential.*

Starting from this previous result, it only takes a few calculations to show that if the mean number of planning dates until t rises linearly with t, then the distribution of these planning dates must be exponential. This follows form collecting the results in steps 3 and 4 since

$$\begin{aligned}
\text{Prob}(V_\tau=t)&=f_1(\tau+t)+\int_0^{\tau} f(u+t)dM(\tau-u)\\
&=\rho(1-\Phi(\tau+t))+\rho\int_0^{\tau} f(u+t)u\\
&=\rho(1-\Phi(t)),
\end{aligned}$$

which holds exactly for all t. But then, since at a planning date $V_{D(i)}$ and $D(i)$ coincide

$$f(t) = \rho(1 - \Phi(t)).$$

This forms a differential equation, with solution

$$f(t) = \rho e^{-\rho t}.$$

The distribution of inattentiveness is exponential. ‖

Proof of Proposition 7. This is the Palm–Khintchine theorem, applied to the set-up in this paper. See Khintchine (1960) for the proof. ‖

Proof of equation (9). The profit function in this model is

$$\pi(p_t(j) - p_t, y_t, a_t) = e^{y_t + (1-\theta)(p_t(j) - p_t)} - e^{y_t - \theta(p_t(j) - p_t) - a}\omega(y_t - \theta(p_t(j) - p_t) - a_t, y_t).$$

In the proof of equation (8), I showed that $y^n - E[y] = -\frac{\pi_{pa}}{\pi_{py}}(a - E[a])$. Using a first-order approximation to $\log \omega(\cdot, \cdot)$ and evaluating the derivatives of the profit function shows that $-\pi_{pa}/\pi_{py} = (1 + \psi)/(\sigma + \psi)$. Similarly, evaluating $-\pi_{py}/\pi_{pp}$ gives the expression for α in the text. To get d^*, just compute π_{pp}/π and use the expression for α. ‖

Acknowledgements. I would like to thank Philippe Aghion, Alberto Alesina, Robert Barro, Susanto Basu, David Laibson, N. Gregory Mankiw, Margaret Meyer, Giuseppe Moscarini, John Muellbauer, Neil Shephard, Lars Svensson, the editor, and two anonymous referees for useful comments; and the Fundação Ciência e Tecnologia, Praxis XXI for financial support.

REFERENCES

AIGINGER, K. (1999), "The Use of Game Theoretical Models for Empirical Industrial Organization", in D. Mueller, A. Haid and J. Weigand (eds.) *Competition, Efficiency and Welfare—Essays in Honor of Manfred Neumann* (Dordrecht: Kluwer Academic Publishers) 253–277.

AKERLOF, G. and YELLEN, J. (1985), "A Near-Rational Model of the Business Cycle with Wage and Price Inertia", *Quarterly Journal of Economics*, **100**, 823–838.

BALKE, N. S. and GORDON, R. J. (1989), "The Estimation of Prewar Gross National Product: Methodology and New Evidence", *Journal of Political Economy*, **97**, 38–92.

BALL, L. and ROMER, D. (1990), "Real Rigidities and the Non-Neutrality of Money", *Review of Economic Studies*, **57**, 183–203.

BARRO, R. J. (1972), "A Theory of Monopolistic Price Adjustment", *Review of Economic Studies*, **39**, 17–26.

BARRO, R. J. and GROSSMAN, H. I. (1971), "A General Disequilibrium Model of Income and Employment", *American Economic Review*, **61**, 82–93.

BASU, S. and FERNALD, J. G. (1997), "Returns to Scale in U.S. Production: Estimates and Implications", *Journal of Political Economy*, **105**, 249–283.

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

BONOMO, M. and CARVALHO, C. V. (2004), "Endogenous Time-Dependent Rules and Inflation Inertia", *Journal of Money, Credit and Banking*, **36**, 1015–1042.

BURSTEIN, A. (2006), "Inflation and Output Dynamics with State Dependent Pricing Decisions", *Journal of Monetary Economics* (forthcoming).

CABALLERO, R. (1989), "Time-Dependent Rules, Aggregate Stickiness and Information Externalities" (Working Paper 428, Columbia University).

CABALLERO, R. and ENGEL, E. M. (1991), "Dynamic (S,s) Economies", *Econometrica*, **59**, 1659–1686.

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

CAPLIN, A. S. and LEAHY, J. (1997), "Aggregation and Optimization with State Dependent Pricing", *Econometrica*, **65**, 601–626.

CARROLL, C. (2003), "Macroeconomic Expectations of Households and Professional Forecasters", *Quarterly Journal of Economics*, **118**, 269–298.

CHARI, V. V., KEHOE, P. J. and MCGRATTAN, E. R. (2000), "Sticky Price Models of the Business Cycle: Can the Contract Multiplier Solve the Persistence Problem?", *Econometrica*, **68**, 1151–1180.

CHEVALIER, J. A., KASHYAP, A. K. and ROSSI, P. E. (2003), "Why Don't Prices Rise During Periods of Peak Demand? Evidence from Scanner Data", *American Economic Review*, **93**, 15–37.

COX, D. (1962) *Renewal Theory* (London: Methuen).

DANZIGER, L. (1999), "A Dynamic Economy with Costly Price Adjustments", *American Economic Review*, **89**, 878–901.

DHYNE, E., ÁLVAREZ, L., BIHAN, H., VERONESE, G., DIAS, D., HOFFMAN, J., JONKER, N., LÜNNEMAN, P., RUMLER, F. and VILMUNEN, J. (2004), "Price Setting in the Euro Area: Some Stylized Facts from Individual Consumer Price Data" (Mimeo, European Central Bank).

DIXIT, A. K. (1991), "Analytical Approximations in Models of Hysteresis", *Review of Economic Studies*, **58**, 141–151.

DOTSEY, M., KING, R. G. and WOLMAN, A. (1999), "State-Dependent Pricing and the General Equilibrium Dynamics of Money and Output", *Quarterly Journal of Economics*, **114**, 655–690.

GOLOSOV, M. and LUCAS, R. E., JR. (2003), "Menu Costs and Phillips Curves" (NBER Working Paper 10187, National Bureau of Economic Research).

KENDRICK, J. W. (1961) *Productivity Trends in the United States* (Princeton, NJ: Princeton University Press).

KHINTCHINE, A. Y. (1960) *Mathematical Methods in the Theory of Queueing*, translated from Russian by D. M. Andrews and M. H. Quenouille (London: Griffin).

KLEMPERER, P. and MEYER, M. (1986), "Price Competition vs. Quantity Competition: The Role of Uncertainty", *Rand Journal of Economics*, **17**, 618–638.

LUCAS, R. E., JR. (1972), "Expectations and the Neutrality of Money", *Journal of Economic Theory*, **4**, 103–124.

LUCAS, R. E., JR. (1976), "Econometric Policy Evaluation: A Critique", *Carnegie-Rochester Series on Public Policy*, **1**, 19–46.

MANKIW, N. G. (1985), "Small Menu Costs and Large Business Cycles: A Macroeconomic Model of Monopoly", *Quarterly Journal of Economics*, **100**, 529–539.

MANKIW, N. G. (2001), "The Inexorable and Mysterious Tradeoff Between Inflation and Unemployment", *Economic Journal*, **111**, C45–C61.

MANKIW, N. G. and REIS, R. (2002), "Sticky Information Sticky Prices: A Proposal to Replace the New Keynesian Phillips Curve", *Quarterly Journal of Economics*, **117**, 1295–1328.

MANKIW, N. G., REIS, R. and WOLFERS, J. (2004), "Disagreement about Inflation Expectations", *NBER Macroeconomics Annual 2003*, **18**, 209–247.

MOSCARINI, G. (2004), "Limited Information Capacity as a Source of Inertia", *Journal of Economic Dynamics and Control*, **28**, 2003–2035.

PENCAVEL, J. (1986), "Labor Supply of Men: A Survey", in O. Ashenfelter and R. Layard (eds.) *Handbook of Labor Economics*, Vol. 1 (Amsterdam: North-Holland) 3–102.

RADNER, R. (1992), "Hierarchy: The Economics of Managing", *Journal of Economic Literature*, **30**, 1382–1415.

REIS, R. (2004), "Inattentive Consumers" (NBER Working Paper 10883, National Bureau of Economic Research).

ROMER, C. D. (1989), "The Prewar Business Cycle Reconsidered: New Estimates of Gross National Product, 1869–1908", *Journal of Political Economy*, **97**, 1–37.

ROSS, S. M. (1983) *Stochastic Processes* (New York: John Wiley).

ROTEMBERG, J. J. (1982), "Monopolistic Price Adjustment and Aggregate Output", *Review of Economic Studies*, **44**, 517–531.

ROTEMBERG, J. J. (2005), "Customer Anger at Price Increases, Changes in the Frequency of Price Adjustments and Monetary Policy", *Journal of Monetary Economics*, **52**, 829–852.

ROTEMBERG, J. J. and WOODFORD, M. (1997), "An Optimization-Based Econometric Framework for the Evaluation of Monetary Policy", *NBER Macroeconomics Annual 1997*, **11**, 297–346.

SHESHINSKI, E. and WEISS, Y. (1977), "Inflation and Costs of Price Adjustment", *Review of Economic Studies*, **44**, 287–303.

SHESHINSKI, E. and WEISS, Y. (1982), "Optimum Pricing Policy Under Stochastic Inflation", *Review of Economic Studies*, **50**, 513–529.

SIMS, C. (2003), "Implications of Rational Inattention", *Journal of Monetary Economics*, **50**, 665–690.

STOKEY, N. and LUCAS, R. E., JR. with PRESCOTT, E. C. (1989) *Recursive Methods in Economic Dynamics* (Cambridge, MA: Harvard University Press).

WEITZMAN, M. L. (1974), "Prices vs. Quantities", *Review of Economic Studies*, **41**, 477–491.

WOODFORD, M. (2003*a*) *Interest and Prices: Foundations of a Theory of Monetary Policy* (Princeton, NJ: Princeton University Press).

WOODFORD, M. (2003*b*), "Imperfect Common Knowledge and the Effects of Monetary Policy", in P. Aghion, R. Frydman, J. Stiglitz and M. Woodford (eds.) *Knowledge, Information, and Expectations in Modern Macroeconomics: In Honor of Edmund S. Phelps* (Princeton, NJ: Princeton University Press) 25–58.

ZBARACKI, M. J., RITSON, M., LEVY, D., DUTTA, S. and BERGEN, M. (2004), "The Managerial and Customer Costs of Price Adjustment: Direct Evidence from Industrial Markets", *Review of Economics and Statistics*, **86**, 514–533.