Inattentive Producers

Ricardo Reis*

Harvard University

This version: June 2004

First draft: April 2004

Abstract

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 behavior. First, I find the conditions under which she prefers to set a plan for the price she charges, or instead prefers to set a plan for the quantity she sells. Second, I show that the agent rationally chooses to be inattentive to news, only sporadically updating her information. I solve for the optimal length of inattentiveness and characterize its determinants. Third, I explicitly aggregate the behavior of many such producers. I apply these results to a model of inflation. I find that the model can fit the post-war data on inflation remarkably well, and the pre-war facts moderately well.

JEL codes: D92, E31, E20

Keywords: Inattentiveness, Production and Pricing under Uncertainty, Inflation, Sticky Information

*I would like to thank Philippe Aghion, Alberto Alesina, Robert Barro, Susanto Basu, David Laibson, N. Gregory Mankiw, Margaret Meyer, John Muellbauer, and Neil Shephard for useful comments; and the Fundacao Ciencia e Tecnologia, Praxis XXI for financial support. Contact: reis@fas.harvard.edu. Address: Department of Economics, Princeton University, Fisher Hall, Princeton, NJ 08544-1021, USA.

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 macroeconomics purports 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. Dotsey, King and Woman (1999), Danziger (1999) and Golosov and Lucas (2003) studied the effects of monetary policy in such an economy. A closely related model of sticky prices has by-passed the micro-foundations and assumed from the start that prices adjust only at some random dates picked from a specific distribution function that allows for simple aggregation (Calvo, 1983, Woodford, 2003a).

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 does not support the model's key micro prediction. Bils and Klenow (2002) noted that individual prices change very frequently in the United States. Finally, many (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. The starting point of these models is the realization 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 world and most of it comes with a cost, in money or time, both in acquiring it but especially also in interpreting it. Following the hallmark of the economic model of 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. 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 behavior, an explicit aggregation across many agents, and 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 behavior 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 price plans at optimally chosen dates. Because the adjustments occur at certain dates regardless of the state of the economy at these dates, the model provides a micro-foundation for time-contingent adjustment. After characterizing the optimal behavior of inattentive agents, I aggregate over many such agents and explicitly characterize the dynamics of aggregate variables in this economy. The inattentive economy exhibits nominal rigidity, since prices only adjust with a delay to shocks. It can therefore be used to study inflation and its relation with real variables, and I do this by contrasting the model's predictions for inflation with data from different periods in U.S. history.

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 (2003) 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 (2002) presents a sticky plan model in which prices also follow pre-determined 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.

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. Finally, Woodford (2003b) and Moscarini (2003) model inattentiveness by price-setters using the alternative Sims' (2003) approach. The model that I use and Sims' model seem to lead to similar predictions. This paper's development of the inattentiveness model should make future comparisons possible. While this is an important research topic, I do not pursue it here.

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 behavior of many inattentive agents. Section 6 uses these results to set up a model of inflation. Sections 7 and 8 contrast the model with data, and section 9 concludes.

2 THE INATTENTIVE PRODUCER'S PROBLEM

2.1 An informal description of the problem

This paper studies the problem of a monopolist producer of a perishable good. Both the production technology and the demand for the 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) made the following insightful observation. Management is essentially about processing information and making decisions. In 1987, 47% of the U.S. workforce was employed in managerial occupations and the number is likely higher today. 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 very substantial. Zbaracki et al. (2003) directly measured the costs incurred by a large U.S. manufacturing firm associated with setting its price catalog. 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 her information sporadically, and to be inattentive to all new information in between adjustment dates. When she does obtain information, 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. This plan of conduct is set conditional on the information at the current planning date.

To illustrate these three decisions, consider the example of a fictional baker. Her first decision is on which variable to write a plan on: price or quantity. If the producer plans to set a price, then she commits to produce whatever amount is required to let the market clear. That is, if she sets a price for her bread, she will keep the oven burning and bread coming out as long as customers are walking through the door.

The baker could instead choose to produce a certain number of breads and give them to a seller. This seller would then take the bread to the market and distribute it among homes and shops. She will charge whatever positive price is necessary to sell all the bread today, since by the end of the day the bread goes bad and becomes worthless. At the end of the day, the seller returns the proceeds from the sales of the bread to the baker.

Facing this option of prices versus quantities, the baker forms an expectation of 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 Walrasian auctioneer that economists routinely assume to ensure that markets clear. While fictional, this auctioneer serves a key role to ensure the consistency of equilibrium economic models. Ideally, one would want to model equilibrium in a way that more explicitly takes into account the information limitations that this paper emphasizes. Barro and Grossman (1971) made important progress in this direction but this remains a challenge for research.

Having decided what to plan, the producer must then decide the content of the plan. If the baker chooses a price plan, this consists of choosing which prices to charge until the next planning date. If she chooses a quantity plan, the content of the plan is a path for the quantity to produce until the next planning date. In both cases, she will make the decision that maximizes expected future profits.

The final decision is when to plan again. When the producer plans today, she decides on the content of the plan as well as on its horizon. When she reaches the end of the optimally chosen planning horizon, she obtains new information and writes a new plan. On the one hand, if the variance of forecast errors grows with time then the longer it has been since the last planning date, the more costly it is to be inattentive since the actions taken under the current plan are severely outdated. On the other hand, by extending the horizon of the plan, the producer saves on the costs of planning. 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 costlessly obtain. For instance, why can't the baker observe the quantity she sold at the end of the day at her fixed price, or hear from the seller at which price did she sell the bread? The answer is that, in principle, she can. But then, the baker must use this one piece of information to infer the current state of demand, and collect a myriad of other pieces of information that affect the consumer's taste for bread or disposable income that will determine demand tomorrow. Moreover, the baker must go through its entire production process and realize how much exactly did she pay for each factor of production 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.

2.2 The formal problem

The monopolist produces a single perishable good with a stochastic technology represented by a continuous and smooth cost function $C(Y, \mathbf{s}) : \mathbb{R}^{S+1} \rightarrow \mathbb{R}$. The quantity produced is denoted by Y and $\mathbf{s}$ 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 $Q(P, \mathbf{s}) : \mathbb{R}^{S+1} \rightarrow \mathbb{R}$, where P stands for the price charged. I assume that

[1] An alternative assumption is that the producer can costlessly acquire a few pieces of information every instant, costlessly absorb these into a sufficient statistic, and costlessly use these to evaluate an optimal plan. 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.

demand is always positive and falls with the price being charged.

The stochastic process $\mathbf{s}_t$ is defined on a standard filtered probability space with filtration $F = \{F_t, t \geq 0\}$. I assume that $\mathbf{s}_t$ 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 $\mathbf{s}_t$ and a set of innovations $u^\tau = (u_t, u_{t+\tau}]$, so that I can write $\mathbf{s}_{t+\tau} = \Psi(\mathbf{s}_{t,} u^\tau)$ to denote the transition between the state at date t and the state at date $t+\tau$.

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, her choices for time t must be conditional on the information she has at time t, which coincides with the available information in the economy at her last planning date.

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

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

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

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

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

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

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

I make the following assumption on this problem:

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

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

i) The maximization problems leading to $\Pi^P(\mathbf{s},t)$ and $\Pi^Y(\mathbf{s},t)$ are well defined for all $\mathbf{s}$ and t; namely, the problems have a solution and expectations can be formed.

ii) $\Pi^P(\mathbf{s},t)$ and $\Pi^Y(\mathbf{s},t)$ are finite for all possible $\mathbf{s}$.

iii) $\Pi(\mathbf{s},t)$ is continuous.

Whenever the agent updates her information and plans, she incurs a non-negative finite cost given by the continuous function $K(\mathbf{s}_t) : \mathbb{R}^S \to \mathbb{R}$. Producers maximize the expected present discounted (at the rate $r > 0$) value of profits including planning costs

$$J(\mathbf{s}_0, D) = E\left\{\sum_{i=0}^{\infty}\left(\int_{D(i)}^{D(i+1)} e^{-rt}\Pi(\mathbf{s}_{D(i)}, t - D(i))dt - e^{-rD(i+1)}K(\mathbf{s}_{D(i+1)})\right)\right\} \tag{3}$$

by choosing a sequence of planning dates $D = \{D(i)\}_{i=0}^{\infty}$ that is $\Im-$measurable.

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

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

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. Bellman's principle of optimality then implies that:[3]

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

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

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 the regulated Brownian motion problems familiar to optimal stopping situations. However, in those problems, the producer observes the

[3]The appendix contains the proof of this and all the other propositions.

state of the economy every instant and decides whether to adjust or not. 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.

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:

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. While being inattentive may be costly, it is equally so wether the plan is set in terms of prices or quantities.

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. Starting with the case of constant marginal costs and considering demand shocks in isolation leads to the following result, where Q_x stands for the partial derivative of the demand function with respect to argument x:

Proposition 3 *With constant marginal costs, up to a second order approximation in the size of the shocks to demand* $\|\mathbf{s}\|$, *plans for prices are preferred if and only if*

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

where all the functions are evaluated at the point where $\mathbf{s} = E[\mathbf{s}]$. *A sufficient, but not necessary, condition is that* $-Q_{ps}P/Q_s \geq 1$.

[4] Weitzman (1974) studies a related but different problem. He 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 commiting ex ante and demand that is moving with shocks.

If the condition holds as an equality, the firm is indifferent between the two types of plans, and if the inequality is reversed, plans for quantities are preferred.

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 degree 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, with price plans rather than quantity plans, positive shocks lead to larger gains and negative shocks to smaller losses. The producer therefore prefers price plans, as stated in the proposition.

Now let the demand function have some curvature ($Q_{pp} \neq 0$). Figure 2 plots the case of an outward shift in demand, but now in the case when $Q_{ps} = 0$ so the slope is unchanged and 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, under a quantity plan price increases only by $\overline{AD}$. Since $\overline{AB} > \overline{AD}$, price plans are preferred. The case of a negative shocks works likewise.

Table 1 evaluates proposition 3 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, 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

Figure 1: Price vs. Quantity Plans – Linear Demand

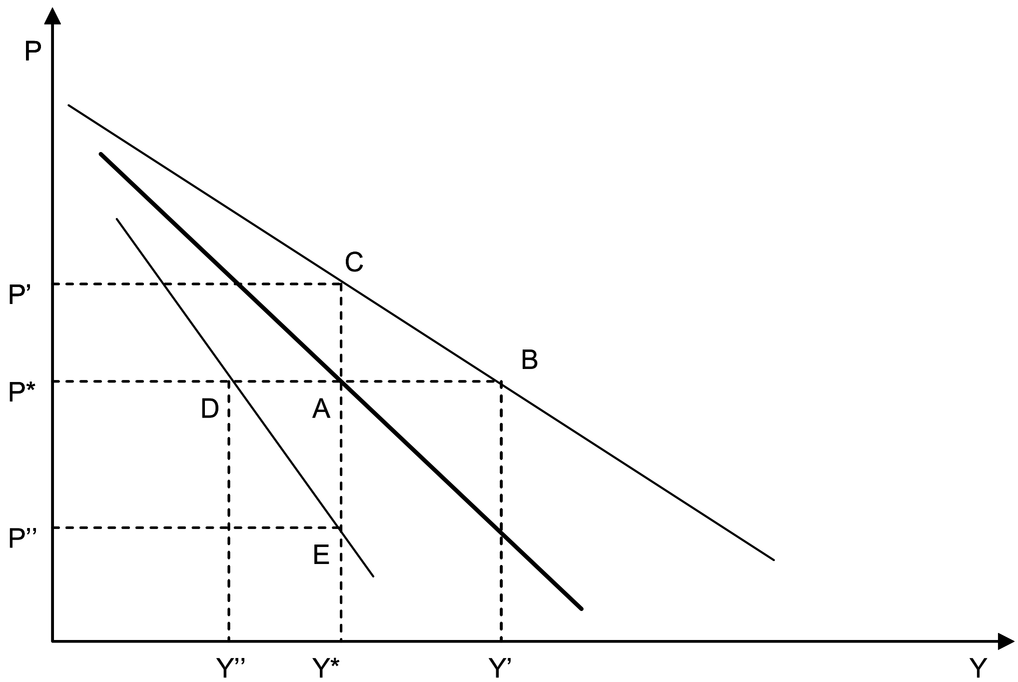
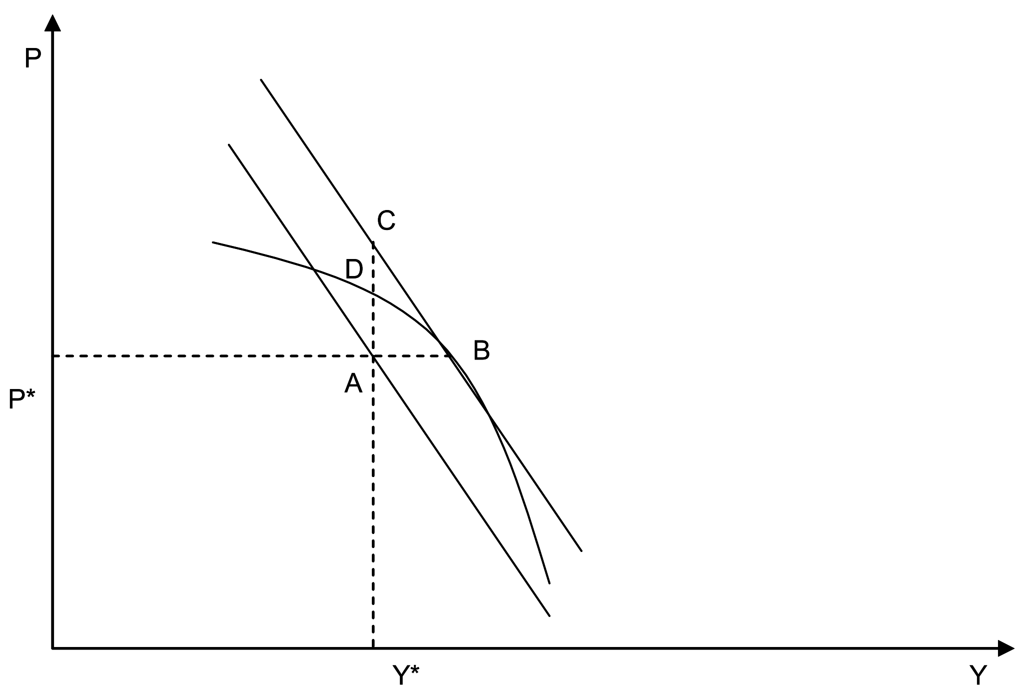


Figure 2: Price vs. Quantity Plans – Concave Demand



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)).

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 $E[s] \leq 2$
			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

Consider now the more general case in which both demand and technology are subject to shocks and are described by arbitrary demand and cost functions. Then:

Proposition 4 *Up to a second order approximation in the size of the shocks to demand* $\|\mathbf{s}\|$*, plans for prices are preferred if and only if*

$$Q_s Q_{ps} + \left(-\frac{Q_s^2}{2Q_p}\right) Q_{pp} + \frac{Q_p^2}{2Q} \left(C_{qq} Q_s^2 + 2C_{qs} Q_s\right) \leq 0 \tag{6}$$

where all the functions are evaluated at the point where $\mathbf{s} = E[\mathbf{s}]$.

There are two new terms in equation (6) involving the slope of marginal costs. The first term shows 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. This effect was emphasized by Klemperer and Meyer (1986) in their study of whether the strategic interaction between firms is best described by the Bertrand or the Cournot models. The second new term 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 lower marginal costs imply that a lower price should be charged. Following a quantity plan is therefore more costly, so a price plan is preferred.

Overall, whether in the world we observe price of quantity plans and whether the choice is as predicted by the model, are interesting questions. The difficulty is to find data on which type of plan firms follow.[5] If such data is available though, the results in this section together with those in Klemperer and Meyer (1986) provide a number of predictions that can be tested.

4 THE DETERMINANTS OF INATTENTIVENESS

4.1 The optimality conditions

Recall that optimal inattentiveness solves:

$$V(\mathbf{s}) = \max_{d} \left\{ \int_0^d e^{-rt} \Pi(\mathbf{s}, t) dt + e^{-rd} E \left[-K(\mathbf{s}_d) + V(\mathbf{s}_d) \right] \right\}$$
$$\text{subject to} : \mathbf{s}_d = \Psi(\mathbf{s}, u^d) \tag{7}$$

The necessary first-order condition for optimality is:

$$\Pi(\mathbf{s}, d) + rE \left[K(\mathbf{s}_d) \right] = E \left[rV(\mathbf{s}_d) + \left(K_s(\mathbf{s}_d) - V_s(\mathbf{s}_d) \right) \frac{\partial \Psi(\mathbf{s}, u^d)}{\partial d} \right] \tag{8}$$

On the left-hand side is the marginal cost of adjusting and on the right-hand side is the marginal benefit; at an optimum the two are equated. The marginal cost equals the sum of two components: the loss in profits earned without adjustment and the flow value of the fixed cost of planning. The marginal benefit has two components as well. The first is the value of being at a planning date. The second captures the changes in planning costs and in the value of being at a planning date that comes with delaying the adjustment. These are the benefits from adjusting today by avoiding larger costs of adjusting tomorrow, and by capturing the future value of having adjusted today rather than tomorrow when this value is expected to be lower.

[5] An exception is Aiginger (1999), who asked a sample of managers of 930 Austrian manufacturing businesses "What is you 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 quanity sold?" In response, 68% of managers professed to follow price plans, while 32% admitted to quantity plans.

Another set of optimality conditions are the envelope theorem conditions with respect to each of the components of the state vector $\mathbf{s}$:

$$V_j(\mathbf{s}) = \int_0^d e^{-rt}\Pi_j(\mathbf{s},t)dt + e^{-rd}E\left[\left(-K_s(\mathbf{s}_d) + V_s(\mathbf{s}_d)\right)\Psi_j(\mathbf{s},u^d)\right]. \tag{9}$$

Equations (7)-(9) characterize the solutions $V(\mathbf{s})$ and $d(\mathbf{s})$ of the optimal inattentiveness problem.

4.2 A general approximate solution

The dynamic program in (7) 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 zero. This approach requires only that $V(\mathbf{s})$ and $d(\mathbf{s})$ are locally differentiable with respect to the costs of planning. Define the function $F(\mathbf{s},t) : R^{S+1} \to R$ as the difference between profits earned with full information and profits earned while following a pre-chosen plan. Then:

Proposition 5 *A perturbation approximation of the optimal inattentiveness around the situation when planning is costless is:*

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

This solution shows that inattentiveness is determined by two factors. First, the larger are the costs of planning, the longer is inattentiveness. Moreover, since $d^*(.)$ 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 behavior differs from optimal behavior 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.

The second determinant of inattentiveness is $F_t(.)$. 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 F_t is if profits are very elastic with respect to price or quantity, so that small errors due to inattentiveness lead to large losses. In these cases, it is more costly to be inattentive, so the agent plans more often.

4.3 The iso-elastic case

One special case is worth solving more 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 demand shock with expectation $\bar{\varepsilon}$, 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.

Using the result in table 1, in this case the producer sets a plan for prices. To maximize profits, she charges:

$$P_t = \frac{\theta}{\theta - 1} E\left[s_t\right]. \tag{10}$$

An interesting property of this demand function is that the optimal price does not depend on shocks to demand. If there were no technology shocks, the producer could be inattentive forever. She does not need to monitor the shocks to demand to set her optimal price.

With shocks to technology, the following result holds:

Proposition 6 *Under the assumptions in this section, optimal inattentiveness is the solution of the equation:*

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

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 increases 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. 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, so the loss from being inattentive is smaller. Thus, the agent stays inattentive for longer.

There is some evidence in favor of this last prediction. Bils and Klenow (2002) 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 United States according to Bils and Klenow (2002). 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 cost of changing prices. That the prices of these 7 goods are adjusted very infrequently is not mysterious. The other three are more interesting: vehicle inspection, legal fees, and safe deposit box rentals. Note that these are goods for which demand is likely not very sensitive to prices, thus supporting the prediction of the model.

4.4 Real and nominal rigidities

More generally, 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 a vector of shocks to productivity, A. The profit function then becomes $\pi(p(j)-p, y, a)$, where small 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 zero so all the producers are attentive. The appendix shows that for an individual producer in an inattentive economy, a second order log-linear approximation of optimal inattentiveness around where the shocks equal their expected value is:[6]

$$d^* = \frac{2}{\alpha}\sqrt{\frac{K}{-\pi_{pp} Var\left[y-y^n\right]}}. \tag{11}$$

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, keeping a price unchanged in response to a shock is close to being optimal. Being inattentive then involves a small cost, so producers are inattentive

[6] If the random variables follow Ito processes, this result is exact.

for longer, precisely as we see in equation (11). Longer inattentiveness in turn implies that prices react with a longer delay to shocks, so having a low α is the key property of the profit function that ensures substantial nominal rigidities. This is the Ball and Romer (1990) result applied to the inattentiveness model.

5 AGGREGATION

5.1 Aggregation with identical firms

In an economy with many inattentive producers, can we say anything in general about how their decisions dates are distributed? 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 very general answers to this question.

Assume that there are many producers in the economy. The sequence of 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 probability that two or more decision dates occur simultaneously for a given producer is zero. This is a statement that the costs of planning are positive almost surely. I also assume that planning dates are not always integral multiples of some non-negative number, so D is not a lattice. That is, I assume that firms do not always stay inattentive for d^* periods, or $2d^*$ periods, or $3d^*$ periods, or so on. 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.

Given that a decision occurred at date 0, let $f_1(t)$ be the probability density function (p.d.f.) for $d(1) = t$. The p.d.f. for how long the next inattentiveness period will last is denoted by $f_2(t)$, and so on, so that $f_i(t)$ is the p.d.f. for the duration of $d(i)$. I assume that:

Assumption 2 *The functions $f_i(t)$ are:*

i) mutually independent;

ii) independent across producers;

iii) the same for all producers.

Independence of decisions dates is convenient since then I only need to keep track of when was the last decision date for each producer. 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.

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 of some firms, due to better organization, management, or economies of scale, having 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 so I leave for future research the task of relaxing these.

It is useful to introduce the following functions:

$F_i(t)$ - the cumulative density function associated with $f_i(t)$,

$G_i(t)$ - the probability that have not planned since date $D(i-1)$, i.e., $G_i(t) = 1 - F_i(t)$,

$A_i(t)$ - the age of a plan, i.e., $A_t = t - D(i-1)$ for $D(i-1) \leq t < D(i)$,

$V_i(t)$ - the remaining duration of the plan, i.e., $V_i(t) = D(i) - t$ for $D(i-1) \leq t < D(i)$,

$I_i(t)$ - the number of plans made by date t, i.e., $I(t) = \{i : D(i) \leq t < D(i+1)\}$,

$H(t)$ - the mean number of planning dates until t, i.e., $H(t) = E[I(t)]$,

ρ - the long-run mean number of planning dates in a unit of time, i.e., $\rho = 1/E[d(i)]$ as $t \to \infty$.

Together, these different functions characterize of the distribution of producers in the economy.

I wish to focus on the properties of an economy that has settled at a steady state after operating for a very long time. Towards this, I introduce the following:

Definition: *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 setup and without any further assumptions, the following remarkable result holds:

Proposition 7 *Under assumption 2, the only stationary equilibrium distribution function for inattentiveness is the exponential distribution with parameter ρ. The distribution $G(t)$ and the distri-*

butions of $A(t)$ and $V(t)$ are all exponential with parameter ρ.

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 is a fortunate result. The exponential distribution is easy to manipulate and its memoryless property allows for tractable aggregation and dynamics.

5.2 Heterogeneous firms

I now drop the requirement that the producers are identical. I still require parts (i) and (ii) of assumption 2, and I further assume that the inattentiveness distribution of each individual agent is stationary.

If each producer has been in operation for a long time, then its distribution of inattentiveness will have converged to the exponential distribution, according to proposition 7. Whatever the characteristics of firm j, proposition 7 states that, with regards to the length of inattentiveness, they can all be summarized by the intensity of attention $\rho(j)$. Then:

Proposition 8 *In an economy with J firms, each with an intensity of attention of $\rho(j)$, and each being at its individual stationary equilibrium distribution of inattentiveness, the distribution of inattentiveness across firms is exponential with parameter $\rho = \sum_{j=1}^{J} \rho(j)$.*

That is, with each firm's planning dates arriving as a Poisson process, the economy's planning dates arrive as a Poisson process as well. This is an exact result that holds even if there are only $J = 2$ firms in the economy.

An even more general result can be established though, by not requiring that each firm has reached its equilibrium distribution. This can be important if there is substantial entry and exit of firms in the economy. At any date, there will be a significant number of firms that have only been in business for a short time and so whose inattentiveness is not well described by proposition 7. In this case, I introduce two new assumptions: (1) that as $J \rightarrow \infty$, then $\sum_{j=1}^{J} \rho(j)$ 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 zero. 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. An application of a famous result known as the Palm-Khintchine theorem then states that:

Proposition 9 *As $J \to \infty$, the distribution of inattentiveness across firms tends to the exponential distribution with parameter ρ.*

The combination of propositions 7, 8 and 9, provide 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 calls 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 interesting because while it is difficult to measure the inattentiveness of economic agents, these other 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 9 is an application. Both mathematics and empirics therefore provide a strong case for exponentially distributed inattentiveness.

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 labor to produce $Y_t(j)$ units of output subject to exogenous stochastic labor productivity A_t. They hire labor in the market paying a real wage $W_t(j)/P_t$. The labor supply function is $\omega(l_t(j), y_t)$, where again small letters denote the logarithms of the respective capital letters. It increases with the amount of labor supplied, with an elasticity of ψ, 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 of profits at the time of planning. This will make it easier to calibrate this parameter.

The profit function then is of the form $\pi(p_t(j) - p_t, y_t, a_t)$ that I studied in section 4.4. A first order log-linearization shows that a firm that last planned at time D charges at time t:

$$p_t(j) = E_D\left[p_t + \alpha(y_t - y_t^n)\right].$$

The natural level of output is the output in the economy if agents are attentive. A first order log-linearization shows that $y_t^n - E[y_t] = \frac{1+\psi}{\sigma+\psi}\left(a_t - E[a_t]\right)$. The parameter α is the Ball and Romer (1990) index of real rigidities, which in this economy equals:

$$\alpha = \frac{\sigma + \psi}{1 + \theta\psi}.$$

Finally, to close the model, I postulate an exogenous stochastic process for nominal income $m_t = p_t + y_t$. This limits the applicability of the model, since it is difficult to think of realistic fiscal or monetary policy shocks as shocks to nominal income. Nevertheless, while monetary policy does not control m_t, it certainly affects it. It is important for monetary policy to understand the relation between nominal income and inflation, even leaving aside the link between the direct policy instruments and nominal income.

More generally, the assumptions that m_t is exogenous and of a given labor supply function $\omega(.)$ allow me to abstract from the consumption and leisure decisions made by inattentive households and on their interaction with inattentive producers. For the purpose of studying inflation and its links to productivity and nominal income, these assumptions are not too restrictive since most existing models share this structure. However, while the model allows for a general study of inflation, it is not adequate to study other macroeconomic variables, such as, for instance, real wages. To do so would require building up the consumption and labor supply sides of the model.[7]

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, table 1 implies that firms will set plans for their prices.

[7] Reis (2003) studies the behavior 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.

The second main theoretical result concerned the optimal choice of the length of inattentiveness. Equation (11) provided an approximation to this choice. Using the profit function for this model, this formula becomes:

$$d^* = \frac{2}{\alpha}\sqrt{\frac{\kappa}{\theta(\theta-1)(\psi+1)Var\left[y_t - y_t^n\right]}}. \tag{12}$$

To assess the predictions of the model for inattentiveness, consider different possible parameter values. For θ, σ and ψ, my preferred parameter values are: $\theta = 10$, since it implies a markup of about 11% which is approximately 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 $\psi = 1/0.15$, to match the estimates of the elasticity of labor supply surveyed in Pencavel (1986). The baseline calibration in Chari, Kehoe and McGrattan (2000) differs only in ψ, since their parameter choices imply that $\psi = 1.25$. Rotemberg and Woodford (1997) estimate using aggregate data that $\theta = 7.88$, $\sigma = 0.16$, and $\psi = 0.47$. Finally, Ball and Romer (1990) set $\sigma = 0$, and use observations on the average markup and the elasticity of labor supply to prefer the values $\theta = 7.8$ and $\psi = 6.7$.

Using the benchmark value of $\sigma = 1$, then up to a constant $y_t^n = a_t = \log(Y_t/L_t)$. I therefore use log output per hour to measure y_t^n. I measure y_t by quarterly real GNP and use an Hodrick-Prescott filter to 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. (2003) 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 catalog. 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. (2003) estimates of the costs of planning as a share of total costs leads to an estimate of 2.8%. I will also consider the impact on inattentiveness of having lower costs of planning, namely 1% and 0.1%.

Table 2 – Optimal expected length of inattentiveness in quarters

		Parameter combinations (θ, σ, ψ)			
		Baseline	Chari et al.	Rotemberg-Woodford	Ball-Romer
		(10, 1, 6.7)	(10, 1, 1.25)	(7.88, 0.16, 0.47)	(7.8, 0, 6.7)
Costs	0.046	10	13	26	12
of	0.028	8	10	20	9
planning	0.010	5	6	12	6
	0.001	2	2	4	2

Table 2 shows the predictions from equation (12) for the average quarters of inattentiveness. A first result to take away from the table is that very small costs of planning can lad 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. (2003) 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 (2003) use data on inflation expectations to infer the speed at which information disseminates in the economy. Both estimate an average inattentiveness of about one 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 an inattentiveness of 4 quarters. Once you take into account the likely magnitude of estimation errors by Zbaracki et al. (2003), these values are plausible. The model is therefore consistent with the separate observations on inattentiveness and 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 her plan, then

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

where $G(j)$ 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. Because there is an infinite number of firms, the third main theoretical result states that $G(j)$ tends to the exponential distribution with parameter ρ which equals the inverse of the average length of inattentiveness. Therefore:

$$p_t = \rho \int_0^\infty e^{-\rho j} E_{t-j} \left[p_t + \alpha(y_t - y_t^n) \right] dj.$$

Taking time derivatives and rearranging, inflation is given by:[8]

$$\dot{p}_t - \alpha \rho p_t = \alpha^2 \rho (y_t - y_t^n) + \rho \int_{-\infty}^t e^{-\rho(t-s)} E_s \left[\dot{p}_t + \alpha(\dot{y}_t - \dot{y}_t^n) \right] ds.$$

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 are always contractionary (although announced disinflations are less contractionary 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.[9]

Mankiw and Reis (2002) reached this Phillips curve by making three assumptions. First, they assumed that agents 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 of producers to such costs. It showed the conditions under which agents choose to set plans for prices. And it found that in a world with many agents, the distribution of inattentiveness converges to the distribution of a Poisson process. The inattentiveness model provides a micro-foundation for the sticky information assumptions.

Having this micro-foundation has many advantages. The model can be used to understand other features of producer behavior aside from pricing, such as for instance the price vs. quantity decision. Moreover, the model provides a unified framework to study different types of behavior 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

[8] I use the standard notation $\dot{x}_t$ to denote the time derivative of a generic variable x_t.

[9] The Calvo (1983) sticky price model, on the other hand, can fit none of these facts.

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 evaluate the impact of different policies.

7 CAN THE MODEL FIT THE POST-WAR FACTS ON INFLATION?

The two relevant reduced-form parameters of the model are α and ρ. Using the baseline parameters for θ, σ and ψ, the implied value of α is 0.11.[10] For ρ, I use the estimates of Carroll (2003) and Mankiw, Reis and Wolfers (2003) and set $\rho = 0.25$ 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. from 1954:1 to 2003:4. Data for the log output per hour in the nonfarm 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.

7.1 Moments and the time-series of inflation

Table 3 uses these parameter values to display the model's predictions for different 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: mean, variability, and persistence. Moreover, it matches well the correlation of inflation with nominal income and productivity, both contemporaneously and with 1-quarter leads and lags. The match is often at the 4^{th} decimal place and, with only one exception, the model's predictions do not differ from the empirical moments by more than 0.05.

[10] 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 (Cooper and John, 1988), so the firms that are adjusting wish to set their individual prices close to those set by non-adjusting firms. Price adjustment is limited, and so monetary shocks have a large real effect. Using the Chari et al (2000) parameters, $\alpha = 0.17$, the Rotemberg and Woodford (1997) estimates, $\alpha = 0.13$, and the Ball and Romer (1990) parameters, $\alpha = 0.13$. Woodford (2003a, 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.

The only significant deviation between the data and the model is that the latter 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 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.

Table 3 – Model vs. data in the post-war U.S.

	Model	Data
Mean(Δp_t)	0.0124	0.0093
St.Dev.(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

Notes: 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.

An alternative way to highlight the model's good fit is to use the empirical innovations to build a series for predicted inflation and compare it with actual inflation. These two series are plotted in figure 3. You can see that the predictions of the model track the data remarkably well. The correlation between predicted and actual inflation is 0.83.

7.2 Impulse responses of inflation

Another test of the model is to examine its predictions for the response of inflation to shocks. The impulse response of inflation to a shock to m_t is particularly important. Policies that affect demand, such as monetary or fiscal shocks, affect inflation through m_t. If the model predicts this impulse response correctly then it should be useful to policymakers in assessing different policies.

Figure 3: Actual and predicted U.S. quarterly inflation

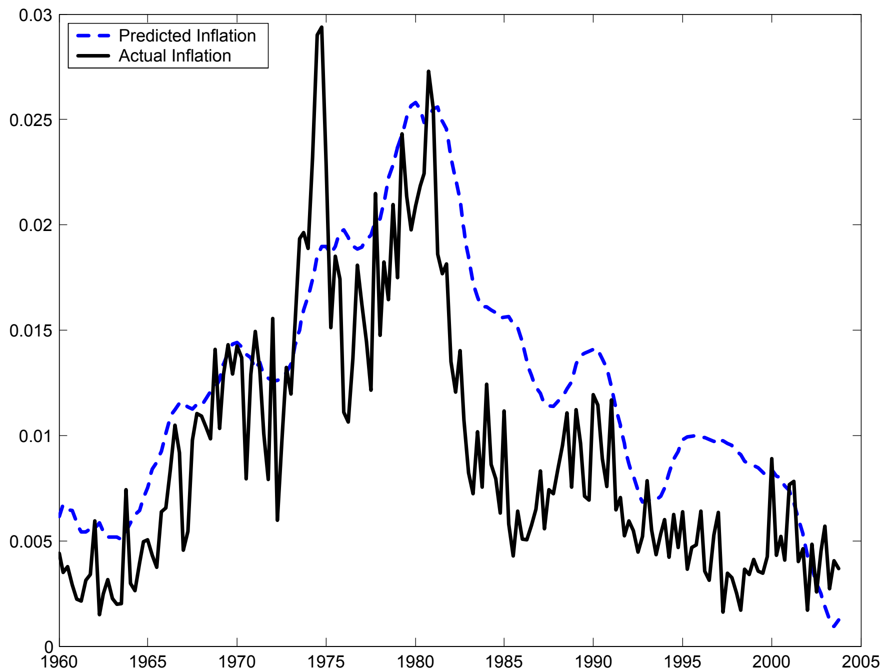
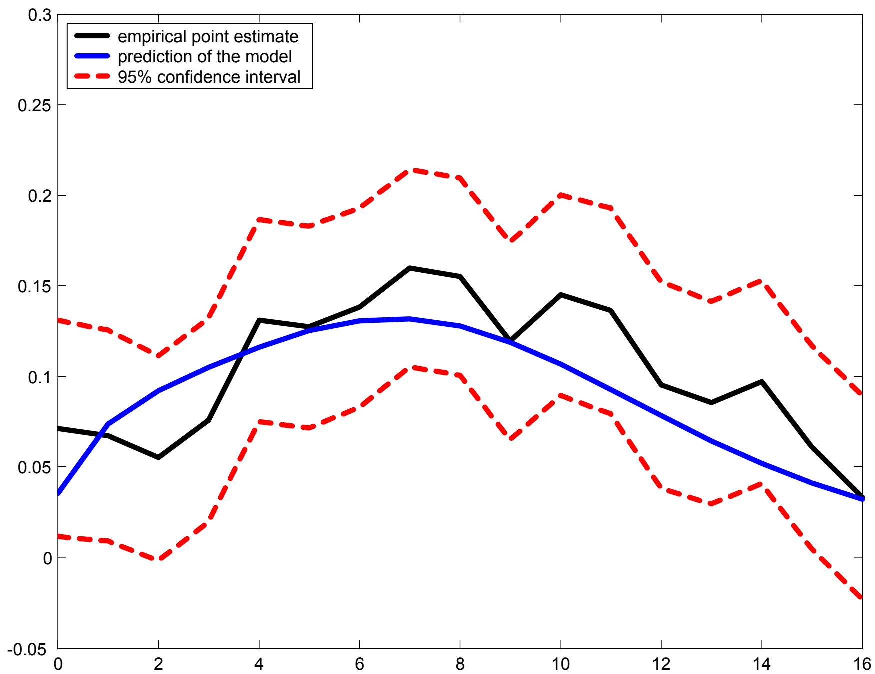


Figure 4 plots the impulse response of inflation to a shock to m_t estimated from the data, together with 95% confidence interval bands.[11] The prediction of the model is also plotted in the figure. Again, the fit is remarkably good. The predicted impulse response is always within the empirical confidence interval and is close to the point estimates. The model is able to capture well the magnitude of the impact of shocks to nominal income on inflation, as well as the fact that the maximum impact occurs 7 quarters after the shock.

Figure 5 plots the impulse response of inflation to shocks to productivity. The model performs slightly worse here, since its predictions differ from the data significantly in the first 2 quarters after the shock. Afterwards though, the model's predictions are within the confidence intervals.

[11] There is an extensive literature estimating such impulse responses. Figure 4 is consistent with the typical reults in this literature (see Christiano, Eichenbaum and Evans, 1999).

Figure 4: Impulse response of inflation to a shock to nominal income



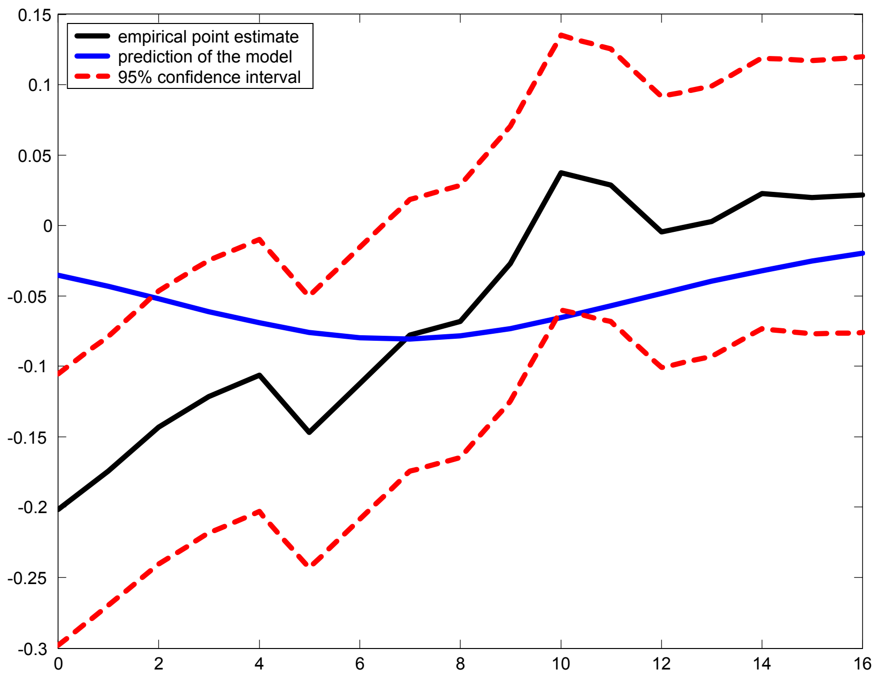
8 CAN THE MODEL FIT THE PRE-WAR FACTS ON INFLATION?

An important use of models is to predict the effect of different policies. This requires that they are structural in that they remain valid across different policy regimes. So far, I have shown that the inattentiveness model can successfully explain inflation in the post-war period. However, during this time, monetary policy did not go through drastic changes. For the most part, policymakers aimed at targeting inflation, even if this goal varied in importance relative to other goals such as stabilizing output.[12] As a result, post-war inflation is well-described by a stable stochastic process.

This was not always the case. 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, imposing a de facto target on the dollar price level. Inflation was close to serially uncorrelated, in stark contrast with the unit-root behavior of the post-war period.

[12] Romer and Romer (2002) give a fascinating account of the changes in the conduct of monetary policy since 1950.

Figure 5: Impulse response of inflation to a shock to productivity



A demanding test of the model is to ask whether it can explain pre-war inflation as well as post-war inflation. This amounts to asking whether the model survives the Lucas (1976) critique.

I obtain data for the pre-war period from three sources. Kendrick (1961) provides estimates of output per hour in the nonfarm sector from 1889 to 1913.[13] There are two different estimates of nominal GNP and its deflator from 1869 to 1913, by Romer (1989) and Balke and Gordon (1989). For the purposes of this paper, the two estimates of nominal income are quite similar. However, the estimates of inflation are very different. I consider both. All of the reliable data for this period comes at an annual frequency.

The stochastic properties of m_t and a_t were markedly different during this period relative to the post-war period. Annual nominal income is now approximately described by a simple random walk, for both the Romer (1989) and the Balke and Gordon (1989) estimates. As for output per

[13] In the beginning of the XXth century, agriculture had a large weight in the U.S. economy with manhours in the farm sector accounting for about 30% of total manhours. Nevertheless, measuring a_t as output per hour in the nonfarm sector or as output per hour in the whole economy is not important for my purposes: the correlation coefficient between the two series is 0.98.

hour, it is approximately white noise in the data.

The fact that the data is annual poses a challenge. Implementing the model requires knowledge of the quarterly processes for m_t and a_t, but there are many different quarterly processes that can generate behavior close to a random walk or to white noise once they are aggregated to an annual frequency. There is no way of using the data to pin down a single statistical model.

I proceed by opting for the most parsimonious statistical representations of nominal income and output per hours that are consistent with the data. For instance, I choose a random walk for quarterly nominal income growth. This implies that annual nominal income should be an IMA(1,1) process, with a moving average coefficient of 0.24. The data does not statistically reject this specification. For productivity, the most parsimonious quarterly process that generates an annual white noise is a quarterly white noise. These statistical models fit the data well. The hope is that the predictions that I will derive are not too sensitive to different assumptions that are also consistent with the data.

Finally, I must take a stand on how the micro parameters have changed between the two time periods and on how this has affected α and ρ. In many ways, the United States after World War II has been quite different from what it was in the beginning of the XXth century. However, there is no clear indication that the elasticity of demand for products or the income and wage elasticities of labor supply were much different then from what they are now. I therefore keep α unchanged at 0.11 in the pre-war period. By keeping this parameter fixed, if I err, I will do so against the model. I will be forcing it to fit two distinct periods with the same parameters.

By the same argument, I will also keep average inattentiveness fixed at 4 quarters. While it is possible that this has changed, it is unclear in which direction. On the one hand, the variability of the output gap fell by 80% between the pre and the post-war periods.[14] In this more volatile pre-war world, agents would wish to plan much more often. On the other hand, it is likely that the great advances in information technology during the XXth century have reduced the costs of planning. Agents would wish to plan less often in the pre-war world, when planning was more costly. To inspect the sensitivity of the results, I also consider an alternative where agents update their plans on average only every three years.

Table 4 contrasts the predictions from the model with the data for the period when I have data

[14] This is true for both the Romer (1989) and the Balke and Gordon's (1989) estimates. While the variability of real GNP is quite different between the two series, the variability of the gap between log real GNP and log output per hour is similar.

on all the series, 1890-1913. The second column of panels A and B have the sample moments in the Romer (1989) and Balke and Gordon (1989) datasets respectively. Note how different these are from the post-war estimates in table 3. It would be remarkable to have a model that could fit both periods.

The second column has the average predictions of the model and the third column has 90% confidence intervals.[15] Table 4 shows that the performance of the model is not quite as successful as in the post-war data. The model substantially over-estimates average inflation and under-estimates the variability of inflation. It also fares poorly in predicting 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. Moreover, the model captures well the dynamic relation between inflation and lagged and lead nominal income. The same is true of the relation between productivity and inflation. The model can therefore match 5 of the 9 moments in the table.

One possible source of bias is the extent of measurement error in the data. This is likely particularly relevant for the measurement of inflation, as the wide differences between Romer (1989) and Balke and Gordon (1989) attest (the correlation between their respective inflation series is only 0.25). Column 3 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. Measurement error can account not just for the discrepancy between the model and the data for the variability of inflation, but also reconciles the predictions with the data for the contemporaneous correlation between inflation and productivity.

Column 5 reports the predictions of the model when agents are inattentive for 3 years on average. The main successes and failures of the model are the same as before.

[15] The estimates in the two panels differ because the volatility of the shocks differs across the two datasets. For instance, nominal income is twice more volatile in the Balke and Gordon (1989) estimates than in Romer's (1989).

Table 4 – Model vs. data in the pre-war U.S.

Panel A – Romer data

	Data	Model	90% Confidence interval	Measurement error adjustment	Model with 3-year updating
Mean(Δp_t)	0.0082	0.0230	0.0180 ; 0.0279	0.0230	0.0230
St.Dev.(Δp_t)	0.0280	0.0085	0.0063 ; 0.0111	0.0280	0.0031
Corr.($\Delta p_t, \Delta p_{t-1}$)	0.1615	0.2242	-0.1985 ; 0.5904	0.0208	0.4910
Corr.($\Delta p_t, \Delta m_t$)	0.7508	0.1184	-0.2020 ; 0.4154	0.0287	-0.0258
Corr.($\Delta p_t, \Delta m_{t-1}$)	0.2482	0.4044	0.0840 ; 0.6601	0.0980	0.0801
Corr.($\Delta p_t, \Delta m_{t+1}$)	0.0187	-0.0758	-0.4056 ; 0.2605	-0.0184	-0.1495
Corr.($\Delta p_t, \Delta a_t$)	-0.0916	-0.6510	-0.8349 ; -0.4347	-0.1577	-0.5311
Corr.($\Delta p_t, \Delta a_{t-1}$)	0.2212	0.3233	0.0746 ; 0.5585	0.0783	0.2636
Corr.($\Delta p_t, \Delta a_{t+1}$)	0.2571	0.3234	0.0729 ; 0.5604	0.0783	0.2622

Panel B – Balke-Gordon data

	Data	Model	90% Confidence interval	Measurement error adjustment	Model with 3-year updating
Mean(Δp_t)	0.0070	0.0229	0.0165 ; 0.0293	0.0229	0.0228
St.Dev.(Δp_t)	0.0285	0.0102	0.0073 ; 0.0137	0.0285	0.0039
Corr.($\Delta p_t, \Delta p_{t-1}$)	0.1225	0.3808	-0.0255 ; 0.6954	0.0487	0.6243
Corr.($\Delta p_t, \Delta m_t$)	0.5343	0.1278	-0.1796 ; 0.4083	0.0354	-0.0306
Corr.($\Delta p_t, \Delta m_{t-1}$)	0.3277	0.4437	0.1537 ; 0.6735	0.1229	0.0885
Corr.($\Delta p_t, \Delta m_{t+1}$)	-0.0960	-0.0853	-0.4074 ; 0.2392	-0.0236	-0.1633
Corr.($\Delta p_t, \Delta a_t$)	-0.0955	-0.5481	-0.7550 ; -0.3207	-0.1518	-0.4365
Corr.($\Delta p_t, \Delta a_{t-1}$)	0.2127	0.2703	0.0265 ; 0.5091	0.0749	0.2172
Corr.($\Delta p_t, \Delta a_{t+1}$)	0.2583	0.2724	0.0272 ; 0.5085	0.0754	0.2175

Notes: The notation Δx_t denotes the annual change in variable x_t. The data column has sample moments for the two data sources for 1890-1913. The predictions of the model come from drawing innovations to nominal income and productivity from normal distributions with variances set to equal the data. The table reports the average over 10,000 simulations, and the 5% and 95% percentiles of the simulated distributions. The second to last column contains the model's predictions under the assumption that classical measurement error accounts for the discrepancy between actual and predicted inflation. The last column has the model's predictions if average inattentiveness is 3 years.

The results in table 4 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. The inattentiveness model not only fits U.S. post-war data remarkably well, but it can also capture many of the dimensions of the data in the pre-war period. Few (if any) of the existing models of inflation would perform this well across such different periods in history.

9 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. For the more commonly used specifications of demand in Economics, these conditions predicted that producers would set plans for prices. Second, I characterized the determinants of the optimally chosen inattentiveness. Third, I showed that under general circumstances, an exponential distribution approximates well the distribution of inattentiveness in an economy with many inattentive agents.

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 took the model to the data and found that it fits the post-war U.S. data on inflation remarkably well. Finally, I asked whether the model could also fit the pre-war facts on inflation. This allowed me to assess whether the model was invariant across policy regimes, and provided a very demanding test of the model. It fared moderately well.

This paper provided the counterpart to the micro-founded sticky-price model of Sheshinski and Weiss (1978, 1982), Caballero and Engel (1991) and Caplin and Leahy (1997), for models of nominal rigidity based on limited information. Combining its modelling tools and results with those in Reis (2003), who studies inattentive consumption choices, provides the foundations for constructing 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.

10 APPENDIX

Proof of Proposition 1: Since $\Pi(\mathbf{s},t)$ and $K(\mathbf{s})$ are well-defined and continuous and D satisfies the measurability restrictions, then $J(\mathbf{s},D)$ is well-defined. From assumption 1, $0 \leq \Pi(\mathbf{s},t) < +\infty$ for all $\mathbf{s}$ and t. The costs of planning are also non-negative and finite. Therefore $J(\mathbf{s},D)$ 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, pages 67-77) then shows that $V(\mathbf{s}) = \max_D J(\mathbf{s},D)$. Since $J(\mathbf{s},D)$ is well-defined and bounded above, so is $V(\mathbf{s})$. The fact that $V(\mathbf{s})$ exists, is unique, and continuous follow from the continuity of $\Pi(\mathbf{s},t)$ and $K(\mathbf{s})$, and the fact that $V(\mathbf{s})$ is the fixed point of a contraction mapping of continuous into continuous functions (Stokey and Lucas, 1989, pages 49-55).■

Proof of Propositions 3 and 4: With full information on the state of the demand shocks $\mathbf{s}$, let the optimal choices of price and quantity be denoted by the functions $P(\mathbf{s})$ and $Y(\mathbf{s})$. These are the solutions from maximizing either $\pi^P(P,\mathbf{s})$ with respect to P, or $\pi^Y(Y,\mathbf{s})$ with respect to Y, respectively. With full information, the solutions to the two problems are of course equivalent: $Y(\mathbf{s}) = Q(P(\mathbf{s}),\mathbf{s})$.

With uncertainty, the optimal price charged in a price plan, P^*, is defined by the first order-condition of the problem in (1): $E(\pi_p^P(P^*,\mathbf{s})) = 0$. A first order Taylor approximation of this equation around $E(\mathbf{s})$ shows that: $P^* = P(E[\mathbf{s}]) + O(\|\hat{\mathbf{s}}\|^2)$. I denote $\mathbf{s} - E[\mathbf{s}]$ by $\hat{\mathbf{s}}$. Using this solution for price in the demand function and doing another first-order Taylor approximation, the quantity sold under a price plan is: $Y^* = Y(E[\mathbf{s}]) + O(\|\hat{\mathbf{s}}\|^2)$. 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 price charged and quantity sold with a quantity-plan are $\hat{P} = P(E[\mathbf{s}]) + O(\|\hat{\mathbf{s}}\|^2)$, and $\hat{Y} = Y(E[\mathbf{s}]) + O(\|\hat{\mathbf{s}}\|^2)$.

Then, note that

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

showing that when $\mathbf{s}$ 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 $\pi^Y(\hat{Y}, E[\mathbf{s}]) = \pi^Y(Y(E[\mathbf{s}]), E[\mathbf{s}]) + O(\|\hat{\mathbf{s}}\|^3)$, so $\pi^P(P^*, E[\mathbf{s}]) - \pi^Y(Y(E[\mathbf{s}]), E[\mathbf{s}])$ is at most third order.

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

$$\pi^P(P^*, \mathbf{s}) - \pi^Y(\hat{Y}, \mathbf{s}) = \pi^P - \pi^Y + \left(\pi_s^P - \pi_s^Y\right)(\mathbf{s} - E[\mathbf{s}]) + \frac{1}{2}\left(\pi_{ss}^P - \pi_{ss}^Y\right)(\mathbf{s} - E[\mathbf{s}])^2 + O(\|\hat{\mathbf{s}}\|^3). \quad (13)$$

All the functions on the right hand side are evaluated at $(P^*, E[\mathbf{s}])$ or $(\hat{Y}, E[\mathbf{s}])$. Consider each of the terms in turn. I already know that $\pi^P - \pi^Y$ is or order $O(\|\hat{\mathbf{s}}\|^3)$. After taking expectations, the second term disappears. As for the third term, since using the definitions of the profit functions, $\pi_{ss}^P = P^* Q_{ss} - C_{qq} Q_s^2 - C_q Q_{ss} - 2C_{qs} Q_s$ and $\pi_{ss}^Y = Q_{ss}^{-1} Q - C_{ss}$, and since $P^* = P(E(\mathbf{s})) + O(\|\hat{\mathbf{s}}\|^2)$, the third term becomes:

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

Finally, use the first-order condition for profit-maximization with respect to prices, $Q + PQ_p = C_q Q_p$, to replace for P. Since price plans are preferred to quantity plans if $\Pi^P(\mathbf{s}, t) > \Pi^Y(\mathbf{s}, t)$, taking expectations of (13), this condition becomes:

$$-\frac{Q}{2Q_p}\left(Q_{ss} + Q_p Q_{ss}^{-1}\right) - \frac{1}{2}\left(C_{qq} Q_s^2 + 2C_{qs} Q_s\right) > O(\|\hat{\mathbf{s}}\|^3).$$

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

$$-\frac{Q}{Q_p^2}\left(Q_s Q_{ps} - \frac{Q_{pp} Q_s^2}{2Q_p}\right) - \left(\frac{C_{qq} Q_s^2}{2} + C_{qs} Q_s\right) > O(\|\hat{\mathbf{s}}\|^3)$$

Rearranging gives the condition in proposition 4.

With constant marginal costs, the term in the second brackets equals zero, so the expression becomes the condition in proposition 3. Moreover, the second-order condition for the price profit-maximization problem is $-Q_{pp}/2Q_p < 1/P$. Combining this inequality with (5) gives the sufficient condition in proposition 3.■

Proof of Proposition 5: Re-write the costs of planning as $K(\mathbf{s}_t) = \kappa^2 \tilde{K}(\mathbf{s}_t)$. I will approximate the solution around the $\kappa = 0$ point.[16] First, subtract the discounted profits obtained from setting prices or quantities with current information on $\mathbf{s}_t$. Using $V(.)$ to denote the value function for this problem (a slight abuse of notation):

$$V(\mathbf{s}) = \max_d \left\{ -\int_0^d e^{-rt} F(\mathbf{s}, t) dt + e^{-rd} E\left[-\kappa^2 \tilde{K}(\Psi(\mathbf{s}, u^d)) + V(\Psi(\mathbf{s}, u^d)) \right] \right\} \tag{14}$$

The optimality conditions are only slightly different:

$$\begin{aligned} -F(\mathbf{s}, d) + r\kappa^2 E\left[\tilde{K}(\mathbf{s}_d)\right] &= E\left[rV(\mathbf{s}_d) + \left(\kappa^2 \tilde{K}_s(\mathbf{s}_d) - V_s(\mathbf{s}_d)\right) \frac{\partial \Psi(\mathbf{s}, u^d)}{\partial d} \right], & (15) \\ V_j(\mathbf{s}) &= -\int_0^d e^{-rt} F_j(\mathbf{s}, t) dt + e^{-rd} E\left[\left(-\kappa^2 \tilde{K}_s(\mathbf{s}_d) + V_s(\mathbf{s}_d)\right) \Psi_j(\mathbf{s}, u^d) \right] & (16) \end{aligned}$$

Since I will be perturbing the solution with respect to κ, the envelope theorem condition with respect to κ is also relevant:

$$V_\kappa(\mathbf{s}) = e^{-rd} E[-2\kappa \tilde{K}(\mathbf{s}_d) + V_\kappa(\mathbf{s}_d)] \tag{17}$$

The system of equations (14)-(17) defines the optimum. When $\kappa = 0$, the solution to the system is $d^* = 0$ and $V(\mathbf{s}) = 0$. At this optimum, $F(\mathbf{s}, 0) = 0$ for all $\mathbf{s}$ and so $F_j(\mathbf{s}, 0) = 0$ as well. Similarly, the nth-order derivatives of V with respect to $\mathbf{s}$ are all zero. Perturbing the system (14)-(17) by differentiating with respect to κ and evaluating at $\kappa = 0$ (where $d^* = 0$, $V = 0$, $V_s = 0$):

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

All the function are evaluated at $\mathbf{s}$ 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:

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

[16] 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 the leading term in an approximation is κ.

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 6: Expected profits under a price plan are:

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

where the second equality follows from the fact that $E(s_d) = s$ for a geometric Brownian motion. The problem in (7) then becomes:

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

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:

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

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

$$A = \max_d \left\{ \frac{\Xi(1 - e^{-rd})}{r} + e^{(b-r)d} \left(-\kappa \Xi + A\right) \right\}. \tag{18}$$

The first-order condition from the maximization problem is:

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

At the optimum d^*, (18) gives the solution for A:

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

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

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

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 (18). Note that:

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

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

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

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

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

Finally, to obtain the approximation, you can use the result in proposition 5. The equation in proposition 6 allows for a check on this result. Let $\tilde{\kappa} = \sqrt{\kappa}$, and note that $H(b, 0, 0) = 0$, that $H_d(b, 0, 0) = 0$ and that $H_{\tilde{\kappa}}(b, 0, 0) = 0$. The implicit function theorem, $H_d d_{\tilde{\kappa}} + H_k = 0$ then does not apply since $H_d = 0$, but because $H_{\tilde{\kappa}} = 0$, the point $\tilde{\kappa} = d = 0$ is a bifurcation point. One further round of differentiation plus the fact that $H_{d\tilde{\kappa}}(b, 0, 0) = 0$ lead to the conclusion that $d_{\tilde{\kappa}} = \sqrt{-H_{\tilde{\kappa}\tilde{\kappa}}/H_{dd}}$. A little more algebra shows that $H_{\tilde{\kappa}\tilde{\kappa}}(b, 0, 0) = -2r(b-r)$ and $H_{dd}(b, 0, 0) = br(b-r)$. Since a first-order Taylor approximation of d^* around $\tilde{\kappa} = 0$ is given by $d^* = d_{\tilde{\kappa}}\sqrt{\kappa}$, the approximation result in the proposition follows.■

Proof of equation (11): If the agent is inattentive, she will set the same price that all other inattentive agents set. Then $p(j) = p$, which solves $E[\pi_p(1, y, a)] = 0$. If she is attentive, she sets price $p(j)^*$ that solves: $\pi_p(p(j)^* - p, y, a) = 0$. A second-order approximation around the point $(1, E[y], E[a])$ of the difference between profits if attentive or inattentive is:

$$\pi(p(j)^* - p, y, a) - \pi(1, y, a) =$$

$$\pi + \pi_p(p(j)^* - p) + \pi_y(y - E[y]) + \pi_a(a - E[a]) + \frac{1}{2}\left[\pi_{pp}(p(j)^* - p)^2 + \pi_{yy}(y - E[y])^2 + \pi_{aa}(a - E[a])^2\right]$$

$$+\pi_{py}(p(j)^* - p)(y - E[y]) + \pi_{pa}(p(j)^* - p)(a - E[a]) + \pi_{ya}(y - E[y])(a - E[a])$$

$$-\pi - \pi_y(y - E[y]) - \pi_a(a - E[a]) - \frac{1}{2}\left[\pi_{yy}(y - E[y])^2 + \pi_{aa}(a - E[a])^2\right] - \pi_{ya}(y - E[y])(a - E[a])$$

All the functions are evaluated at $(1, E[y], E[a])$. Cancelling common terms and since $\pi_p = 0$, gives:

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

The natural level of output is defined by $\pi_p(1, y^n, a) = 0$; 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:

$$\pi(p(i)^* - p, y, a) - \pi(1, y, a) = -\frac{\pi_{pp}\alpha^2(y - y^n)^2}{2}$$

From the definition of the $F(\mathbf{s}, t)$ function, it then follows:

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

Since $Var[y - y^n] = E[(y - y^n)^2] - (E[y] - E[y^n])^2$, and since the equation defining the natural level of output implies that $E[y^n] - E[y]$ is second order, it follows that:

$$F(\mathbf{s}, t) = -\frac{\pi_{pp}\alpha^2 Var[y_t - y_t^n]}{2}$$

Finally, $F_t(\mathbf{s}, 0)$ 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.

Proof of Proposition 7: This proof proceeds over a sequence of steps.

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

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 $h(\tau, t)/F(\tau)$. As $\tau \to 0$, this is the conditional probability that no decision dates occur in period t, under the condition that the first moment of this period was a decision date. This is Palm's function: $\varphi(t) = \lim_{\tau \to 0} h(\tau, t)/F(\tau)$. Khintchine (1960, pages 45-48) proves the following result:

Theorem: $F_k(t) = 1 - \varphi(t)$ for all $k \geq 2$.

I therefore only need to describe two distributions, $F_1(t)$, and $F(t) = F_k(t)$ for all $k \geq 2$.

Step 2: *Proving the following result*:

Theorem (Wald): $E\left[\sum_{i=1}^{I} d(i)\right] = E[I]E[d(i)]$

Proof: Let $\iota(i) = 1$ if $i \leq I$ but $\iota(i) = 0$ if $i > I$. Then, $\sum_{i=1}^{I} d(i) = \sum_{i=1}^{\infty} d(i)\iota(i)$ so:

$$E\left[\sum_{i=1}^{I} d(i)\right] = E\left[\sum_{i=1}^{\infty} d(i)\iota(i)\right] = \sum_{i=1}^{\infty} E\left[d(i)\iota(i)\right],$$

where the last equality uses the independence of the $d(i)$. But then, note that $\iota(i)$ is determined by the first $I - 1$ decision dates and so is independent of the occurrence of the next decision date $d(I)$. It then follows that:

$$\sum_{i=1}^{\infty} E\left[d(i)\iota(i)\right] = E\left[d(i)\right] \sum_{i=1}^{\infty} E\left[\iota(i)\right] = E\left[d(i)\right] E\left[I\right]$$

Step 3: *Proving the following result:*

Elementary Renewal Theorem: $\rho = \lim_{t \to \infty} H(t)/t$.

Proof: From the definition of $D(i)$ and $I(t)$, is follows that:

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

Taking expectations and using Wald's theorem from step 2:

$$\begin{aligned} E\left[d(i)\right] E\left[I(t)+1\right] &> t \Leftrightarrow \\ \frac{H(t)+1}{t} &> \frac{1}{E\left[d(i)\right]}. \end{aligned}$$

The second line follows from the definition of $H(t)$ and re-arranging. Taking the limit:

$$\liminf_{t\to\infty} \frac{H(t)}{t} \geq \frac{1}{E\left[d(i)\right]}.$$

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 $i = 1, 2, ...$ This in turn defines, $\bar{D}(i) = \sum_{j=1}^{i} \bar{d}(j)$, $\bar{I}(t) = \sup\left\{i : \bar{D}(i) \leq t\right\}$, and $\bar{H}(t) = E[\bar{I}(t)]$. 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:

$$\limsup_{t\to\infty} \frac{\bar{H}(t)}{t} \leq \frac{1}{E\left[\bar{d}(i)\right]}.$$

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

$$\limsup_{t\to\infty} \frac{H(t)}{t} \leq \frac{1}{E\left[d(i)\right]}.$$

I have then shown that:

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

This proof assumed that $E\left[d(i)\right] < \infty$. If $E\left[d(i)\right] = \infty$ a similar proof holds using the truncated process and since $E\left[\bar{d}(i)\right] \to \infty$, now $\rho \to 0$.

Step 4: Finding the distribution $f_1(t)$.

From the definition of V_t, the time until the next planning date:

$$Prob(V_\tau = t) = f_1(\tau + t) + \int_0^\tau f(t+u)dH(\tau - u).$$

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 zero (which I will verify later). Then, by the elementary renewal theorem:

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

The second equality uses the definition of $F(t)$. It is important to note the following:

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

does not hold generally under only the elementary renewal theorem but requires a closely related alternative called the Key Renewal Theorem (Ross, 1983, pages 61-65). Under many cases though, the elementary renewal theorem suffices.

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

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

Step 5: Proving that $H(t) = \rho t$.

Using the definition of $H(t)$:

$$\begin{aligned} H(t) &= \sum_{i=1}^{\infty} i Prob[I(t) = i] \\ &= \sum_{i=1}^{\infty} i\,(Prob[D(i) \leq t] - Prob[D(i+1) \leq t]) \end{aligned}$$

But, $Prob[D(i) \leq t] = F_1 * F_{i-1}(t)$ where $*$ stands for a convolution. Then:

$$
\begin{aligned}
H(t) &= \sum_{i=1}^{\infty} i\,(F_1 * F_{i-1}(t) - F_1 * F_i(t)) \\
&= F_1 + \sum_{i=1}^{\infty}(i+1)F_1 * F_i(t) - \sum_{i=1}^{\infty} iF_1 * F_i(t) \\
&= \sum_{i=1}^{\infty} F_1 * F_{i-1}(t)
\end{aligned}
$$

The Laplace transform of it is (using the fact that $F_i(t) = F(t)$ for $i \geq 2$ from step 1):

$$
\mathcal{L}(H(s)) = \frac{\mathcal{L}(F_1(s))}{1 - \mathcal{L}(F(s))}
$$

The Laplace transform of the initial distribution is:

$$
\begin{aligned}
\mathcal{L}(F_1(s)) &= \frac{\mathcal{L}(f_1(s))}{s} \\
&= \frac{\mathcal{L}(\rho(1 - F(s)))}{s} \\
&= \frac{\rho\,(1 - \mathcal{L}\,(F(s)))}{s}
\end{aligned}
$$

where the first and third equalities are standard results for Laplace transforms, and the second equality comes from using the result in step 4. Going back to $\mathcal{L}(H(s))$ and substituting for $\mathcal{L}(F_1(s))$:

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

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

Step 6: Proving that the distribution is exponential.

Collecting the results in steps 4 and 5, then:

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

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

$$f(t) = \rho(1 - F(t))$$

This forms a differential equation, with solution:

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

The distribution of inattentiveness is exponential. That the other distributions in the proposition are also exponential follows easily.■

Proof of Proposition 8: The process of arrival of decision dates for each firm is an independent Poisson process. An elementary result is that the sum of J independent Poisson processes with parameter $\rho(j)$ is a Poisson process with rate $\sum_{j=1}^{J} \rho(j)$.■

Proof of Proposition 9: This just states the Palm-Khintchine theorem, applied to the setup in this paper. See Khintchine (1960) for the proof.■

Results for the model in section 6: The profit function in this model is:

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

In the proof of equation (11) I showed that $y^n - E[y] = -\frac{\pi_{pa}}{\pi_{py}}(a - E[a])$. Using a first-order approximation to $\log \omega(.,.)$ and evaluating the derivatives of the profit function shows that $-\pi_{py}/\pi_{pp} = (1+\psi)/(\sigma+\psi)$. 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.■

References

[1] Aiginger, Karl (1999) "The Use of Game Theoretical Models for Empirical Industrial Organization," in D. C. Mueller, A. Haid, and J. Weigand, eds., *Competition, Efficiency and Welfare - Essays in Honor of Manfred Neumann*, Kluwer Academic Publishers, Dordrecht.

[2] Akerlof, George and Janet Yellen (1985) "A Near-Rational Model of the Business Cycle with Wage and Price Inertia," *Quarterly Journal of Economics*, vol. 100, Supplement, pp. 823-838.

[3] Balke, Nathan S. and Robert J. Gordon (1989) "The Estimation of Prewar Gross National Product: Methodology and New Evidence," *Journal of Political Economy*, vol. 97, pp. 38-92.

[4] Ball, Laurence and David Romer (1990) "Real Rigidities and the Non-Neutrality of Money," *Review of Economic Studies*, vol. 57, pp. 183-203.

[5] Barro, Robert J. (1972) "A Theory of Monopolistic Price Adjustment," *Review of Economic Studies*, vol. 39, pp. 17-26.

[6] Robert J. Barro and Herschel I. Grossman (1971) "A General Disequilibrium Model of Income and Employment," *American Economic Review*, vol. 61, pp. 82-93.

[7] Basu, Susanto and John G. Fernald (1997) "Returns to Scale in U.S. Production: Estimates and Implications," *Journal of Political Economy*, vol. 105, pp. 249-283.

[8] Bils, Mark and Peter J. Klenow (2002) "Some Evidence on the Importance of Sticky Prices," NBER working paper no. 9069.

[9] Bonomo, Marco and Carlos Viana de Carvalho (2003) "Endogenous Time-Dependent Rules and Inflation Inertia," Princeton University, unpublished.

[10] Burstein, Ariel (2002) "Inflation and Output Dynamics with State Dependent Pricing Decisions," University of Michigan, unpublished.

[11] Caballero, Ricardo (1989) "Time-Dependent Rules, Aggregate Stickiness and Information Externalities," Columbia University working paper no. 428.

[12] Caballero, Ricardo and Eduardo Engel (1991) "Dynamic (S,s) Economies," *Econometrica*, vol. 59 (6), pp. 1659-1686.

[13] Calvo, Guillermo A. (1983) "Staggered Prices in a Utility Maximizing Framework," *Journal of Monetary Economics*, vol. 12, pp. 383-398.

[14] Caplin, Andrew S. and John Leahy (1997) "Aggregation and Optimization with State Dependent Pricing," *Econometrica*, vol. 65, pp. 601-626.

[15] Carroll, Christopher (2003) "Macroeconomic Expectations of Households and Professional Forecasters," *Quarterly Journal of Economics*, vol. 118, pp. 269-298.

[16] Chari, V. V., Patrick J. Kehoe, and Ellen R. McGrattan (2000) "Sticky Price Models of the Business Cycle: Can the Contract Multiplier Solve the Persistence Problem?" *Econometrica*, vol. 68, pp. 1151-1180.

[17] Christiano, Lawrence J., Martin Eichenbaum, and Charles L. Evans, (1999) "Monetary Policy Shocks: What Have We Learned and To What End?" *Handbook of Macroeconomics*, eds. J. B. Taylor and M. Woodford, pp. 65-148, Elsevier, Amsterdam.

[18] Cooper, Russell and Andrew John (1988) "Coordinating Coordination Failures in Keynesian Models," *Quarterly Journal of Economics*, vol. 103, pp. 441-463.

[19] Cox, David R. (1962) *Renewal Theory*, Methuen, London.

[20] Danziger, Leif (1999) "A Dynamic Economy with Costly Price Adjustments," *American Economic Review*, vol. 89, pp. 878-901.

[21] Dotsey, Michael, Robert G. King, and Alexander Wolman (1999) "State-Dependent Pricing and the General Equilibrium Dynamics of Money and Output," *Quarterly Journal of Economics*, vol. 114, pp. 655-690.

[22] Golosov, Mikhail and Robert E. Lucas Jr. (2003) "Menu Costs and Phillips Curves," NBER working paper no. 10187.

[23] Kendrick, John W. (1961) *Productivity Trends in the United States*, Princeton University Press, Princeton.

[24] Paul Klemperer and Margaret Meyer (1986) "Price Competition vs. Quantity Competition: The Role of Uncertainty," *Rand Journal of Economics*, vol. 17, pp.618-638.

[25] Khintchine, Aleksandr Y. (1960) *Mathematical Methods in the Theory of Queueing*, translated from Russian by D. M. Andrews and M. H. Quenouille, Griffin, London.

[26] Lucas, Robert E. Jr. (1972) "Expectations and the Neutrality of Money," *Journal of Economic Theory*, vol. 4, pp. 103-124.

[27] Lucas, Robert E. Jr. (1976) "Econometric Policy Evaluation: A Critique," *Carnegie-Rochester Series on Public Policy*, vol. 1, pp. 19-46.

[28] Mankiw, N. Gregory (1985) "Small Menu Costs and Large Business Cycles: A Macroeconomic Model of Monopoly," *Quarterly Journal of Economics*, vol. 100, pp. 529-539.

[29] Mankiw, N. Gregory (2001) "The Inexorable and Mysterious Tradeoff Between Inflation and Unemployment," *Economic Journal*, vol. 111, pp. C45-C61.

[30] Mankiw, N. Gregory and Ricardo Reis (2002) "Sticky Information Versus Sticky Prices: A Proposal to Replace the New Keynesian Phillips Curve," *Quarterly Journal of Economics*, vol. 117, pp. 1295-1328.

[31] Mankiw, N. Gregory, Ricardo Reis and Justin Wolfers (2003) "Disagreement about Inflation Expectations," *NBER Macroeconomics Annual 2003*, forthcoming.

[32] Moscarini, Giuseppe (2003) "Limited Information Capacity as a Source of Inertia," *Journal of Economic Dynamics and Control*, forthcoming.

[33] Pencavel, John (1986) "Labor Supply of Men: A Survey," in *Handbook of Labor Economics*, vol. 1, O. Ashenfelter and R. Layard, eds., North-Holland, Amsterdam.

[34] Radner, Roy (1992) "Hierarchy: The Economics of Managing," *Journal of Economic Literature*, vol. 30, 1382-1415.

[35] Reis, Ricardo (2003) "Inattentive Consumers," Harvard University, unpublished.

[36] Romer, Christina D. (1989) "The Prewar Business Cycle Reconsidered: New Estimates of Gross National Product, 1869-1908," *Journal of Political Economy*, vol. 97, pp. 1-37.

[37] Romer, Christina D. and David Romer (2002) "The Evolution of Economic Understanding and Postwar Stabilization Policy," in *Rethinking Stabilization Policy*, Federal Reserve Bank of Kansas City.

[38] Ross, Sheldon M. (1983) *Stochastic Processes*, John Wiley and Sons, New York

[39] Rotemberg, Julio (1982) "Monopolistic Price Adjustment and Aggregate Output," *Review of Economic Studies*, vol. 44, pp. 517-531.

[40] Rotemberg, Julio and Michael Wodford (1997) "An Optimization-Based Econometric Framework for the Evaluation of Monetary Policy," *NBER Macroeconomics Annual 1997*, pp. 297-346.

[41] Sheshinski, Eytan and Yoram Weiss (1977) "Inflation and Costs of Price Adjustment," *Review of Economic Studies*, vol. 54, pp. 287-303.

[42] Sheshinski, Eytan and Yoram Weiss (1982) "Optimum Pricing Policy Under Stochastic Inflation," *Review of Economic Studies*, vol. 50, pp. 513-529.

[43] Sims, Christopher (2003) "Implications of Rational Inattention," *Journal of Monetary Economics*, vol. 50, pp. 665-690.

[44] Stokey, Nancy and Robert E. Lucas Jr., with Edward C. Prescott (1989) *Recursive Methods in Economic Dynamics*, Harvard University Press, Cambridge.

[45] Weitzman, Martin L. (1974) "Prices vs. Quantities," *Review of Economic Studies*, vol. 41, pp. 477-491.

[46] Woodford, Michael (2003a) *Interest and Prices: Foundations of a Theory of Monetary Policy*, Princeton University Press, Princeton.

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

[48] Zbraracki, Mark J., Mark Ritson, Daniel Levy, Shantanu Dutta, and Mark Bergen (2003) "The Managerial and Customer Costs of Price Adjustment: Direct Evidence from Industrial Markets," *Review of Economics and Statistics*, forthcoming.