Rigid Pricing and Rationally Inattentive Consumer*

Filip Matějka

PACM, Princeton University and CERGE-EI[†]

WP 409

Abstract

This paper proposes a mechanism leading to rigid pricing as an optimal strategy. It applies a framework of rational inattention to study the pricing strategies of a monopolistic seller facing a consumer with limited information capacity. The consumer needs to process information about prices, while the seller is perfectly attentive. It turns out that the seller chooses to price discretely even for a continuous range of unit input costs, i.e. charges a finite set of different prices only. The price usually stays constant when unit input cost changes only a little. The seller does so to provide the consumer with easily observable prices and thus stimulate her to consume more. In the model's dynamic version, this mechanism implies that prices respond to cost shocks with a delay.

Abstrakt

Tento článek navrhuje mechanizmus generující rigidní ceny jako optimální strategii. Používá teorii racionální nepozornosti ke studiu cenových strategií monopolistického prodejce, jehož zákazníci mají omezenou informační kapacitu. Spotřebitelé musí informace o cenách zpracovávat, zatímco prodejce má všechny informace přesné. Ukazuje se, že prodejce schválně vybírá pouze z několika různých cen, a to i pokud se jeho vstupní náklady pohybují ve spojitém intervalu. Cena pak obvykle zůstává konstantní pokud se vstupní náklady změní pouze o trochu. Prodejce volí tuto strategii, aby byly ceny pro spotřebitele lépe pozorovatelné, a ti potom spotřebovávají více. V dynamické verzi modelu tento mechanizmus způsobuje zpožděné reakce cen na šoky.

---

1 Introduction

This paper studies the implications of consumers' inattentiveness to sellers' choices of pricing strategies. The framework of rational inattention - which was introduced by Sims in [10], [11] and [12] - assumes that agents have limited information capacity. The presented model allows for full flexibility of information processing[3].

Consumers often do not realize what a product's exact price is at the moment of a purchase decision. However, they certainly have some knowledge about the price and pay more attention to some prices than they do to others. Some consumers often grab certain products in a supermarket without even looking at their prices. Many of us at least read price's first few digits, while ignoring the cents. Typically, we implicitly assume that prices end with .95 or .99. If the number of cents is actually 85, we may not spot it and still keep our initial guess. Sometimes, we read just the first digit only or none at all.

If it is unpleasant, i.e. costly, to inspect prices and if uncertainty about the true price can discourage consumption, then sellers could try to accommodate consumers with more predictable prices. It might be optimal for the seller not to respond to every minor change of input cost. Such frequent price changes would require consumers to pay lots of attention to the price, and if they did not want to, then they could rather decide to consume less.

In the model, which is presented in the following sections, the consumer has limited information capacity. She processes information about a price set by a monopolistic seller and decides how much of the seller's product to purchase. The seller chooses his pricing strategy in advance and commits to it. The strategy defines what price is charged at each particular unit input cost. The pricing strategy together with the distribution of unit input costs then determine the distribution of prices. The consumer knows the distribution of prices; it forms her prior knowledge about prices.

When the seller designs a pricing strategy, he realizes that the consumer has limited information capacity and that the strategy forms the consumer's prior. The more dispersed the pricing is, the more difficult it is for the consumer to

---

[3]No specific constraints on the shape of received signals, such as gaussianity, are imposed. For more details of such a setup see Matějka [6]

observe the true realized price. If the consumer has a specific preferences and is uncertain about the true value of price, then he might relocate a portion of her spending towards other products. This behavior is an analog to precautionary savings. Therefore, there is a preference for the distribution of prices to be rather concentrated. In turns out that the seller often chooses a few finite subintervals of the range of unit input costs and charges one price in each of the subintervals only. The seller benefits from a consumer having better knowledge about prices, which makes her consume more.

After analyzing the static model, the paper continues with a discussion of the potential implications of this mechanism for dynamic models of pricing. Taking some standard properties of knowledge refinement through a limited capacity channel as given, it is shown that it is optimal for sellers to respond to aggregate shocks sluggishly. The idea is that while consumers gradually learn about the nature of a shock, they refine their knowledge about the seller's potential responses to it. The seller desires to price in line with consumers' expectations and thus, the slower consumers learn, the slower he changes the price. By doing so, consumers' knowledge about the price is more precise and consumption thus increases.

Assessing nominal rigidities is at the heart of New Keynesian economics. Traditional explanations of nominal rigidities are based on explicit costly adjustment of prices (menu cost models) or even on a complete inability to alter prices between two a priori given moments (Calvo models). More recently, there has been lots of interest in the nominal rigidities implied by the seller's inability to process or acquire perfect information. Mackowiak and Wiederholt[4], Woodford[14] and Matějka[6], study pricing of rationally inattentive sellers. Reis[9] adopts a sticky information approach, which was introduced in Mankiw and Reis[5]. All of these models have frictions microfounded on the sellers' side.

This paper, on the other hand, analyzes the implications of consumers' inattentiveness for patterns of pricing. In reality, consumers can often possess less information about prices than sellers about their input costs - especially in markets where sellers offer just one product, but consumers need to compare prices of several different sellers. The proposed mechanism resembles the explanation of nominal rigidities via the existence of implicit contracts between sellers

3

and consumers[4], which are related to consumers' preferences for stable nominal prices. Blinder[1], and also Fabiani et al.[3], found through a series of interviews with price-setters that implicit contracts are likely to be one of the prominent reasons for prices to be sticky.

2   The Model

A rationally inattentive consumer interacts with a monopolistic seller. The consumer processes information about the price, p, of the seller's product and then decides how much of the good to consume, c. The seller incurs a stochastic unit input cost, μ, and applies his pre-selected pricing strategy to determine what price to charge. At the chosen price, the seller provides whatever amount the consumer wishes to buy.

Information processing generates a transformation of prior knowledge into posterior knowledge. Prior knowledge is what the consumer knows before she starts processing information. The consumer is rational, the probability distribution of prices representing her prior knowledge coincides with the true distribution of prices, which is given by the seller's pricing strategy.

In other papers on rational inattention, Sims [12] and Mateějka [6], the prior was fully determined by a pdf g. We will however be concerned with cases where the price distribution has lumps of probability located at single points; such distributions do not have a pdf with respect to a Lebesgue measure. We will therefore describe the true distribution of prices by its probability measure ω_p; a pdf of prices with respect to this measure is then $g_T(p) = 1$. Prior knowledge has a pdf $g_K(p)$ with respect to ω_p. By assuming that the consumer is rational, we require $g_K(p) = 1$.

The consumer derives utility from consuming the seller's product and a basket of other goods. After purchasing from the seller, she automatically spends all that is left of the initial nominal endowment, e, on purchasing the basket. Let both the endowment and the basket's price be known and equal to 1. Let the indirect utility from consuming an amount c at a price p be denoted by $U(p, c)$. In this paper, the utility $U(p, c)$ will mostly take the form of either a

---

[4]See Nakamura and Steinsson[8]. They provide an explanation for the coexistence of a rigid regular price and frequent deviations from it based on consumers' habits in single goods.

CES aggregator, or its limit version.

$$U_1(p, c) = \left(a \cdot c^r + (1 - a) \cdot (1 - pc)^r \right)^{1/r}, \tag{1}$$

where $a \in (0, 1)$ is a share parameter and $r = 1 - 1/\theta$ is the elasticity of substitution, $\theta \in (1, \infty)$. A scaled aggregator's limit as the share parameter a approaches zero is:

$$U_2(p, c) = c^{1-1/\theta} - pc. \tag{2}$$

A rationally inattentive consumer chooses what pieces of information to process. Then, she decides how much to consume based on her posterior knowledge about price. Given the distribution of prices, the optimal pieces of information together with optimal consumption responses form a **consumer's strategy**.

The seller in turn considers all consumer's strategies as responses to his pricing strategies and selects the optimal pricing to maximize his expected profit,

$$E[\Pi] = E[c \cdot (p - \mu)]. \tag{3}$$

A **seller's pricing strategy** is a rule determining what price to charge at an incurred unit input cost. Given a distribution of input costs, a pricing strategy determines a distribution of prices and thus the consumer's prior knowledge. The seller moves first. In equilibrium, a pricing strategy maximizes the seller's profit, while the consumer's strategy is an optimal response to a price distribution determined by the seller.

The timing of events is as follows:

1. The seller picks his pricing strategy once and for all.

2. The chosen pricing strategy determines the distribution of prices and forms the consumer's prior knowledge.

3. The consumer picks her strategy of information processing and consumption responses.

4. The seller's unit input cost is drawn by nature; price and consumption are executed according to the pre-selected pricing and consumption strategies.

2.1 Inattentive Consumer

The consumer is rationally inattentive. She chooses how much of the seller's product to purchase, while everything that is left of her initial endowment is automatically used to purchase a basket of other goods.

Since the consumer is inattentive, she needs to processes information about the actual price before deciding on the consumption amount. She starts with some prior knowledge about price. After she has processed new information, her posterior knowledge is given by another distribution that is typically more concentrated.

Under rational inattention, the agent's ability to process information is limited, but she is still completely free to process the pieces of information she cares about the most. This model takes full advantage of the flexibility of rational inattention, which provides a rigorous treatment of information processing with respect to different forms of the prior. The prior coincides with the true distribution of prices,

$$g_K(P) = g_T(P) = 1. \tag{4}$$

In general, the consumer chooses what pieces of information about price to process based on i)What she knew in advance, ii)the relative importance of various pieces of information - given by the form of her utility function, iii) The actual realized price, and iv)the realization of a noise component.

The consumer's decision strategy is fully described by a joint distribution of p and c, given by a pdf $f(P, C)$ and probability measures ω_p and ω_c. They define both the choices of signals and the choices of consumption responses to realized posteriors. A conditional distribution of price P given c represents a posterior about P leading to a selected amount of consumption equal to c, while a $f(C|p)$ together with ω_c determine a distribution of consumption given a realized price, p. The consumers is given a marginal distribution of p. What she is left to choose is the collection of conditional distributions of consumption c.

Definition 1 *The consumer's problem: Let $U(P, C)$ be the indirect utility function and κ be the consumer's information capacity. A pdf $g_K(P)$ with respect to a given probability measure of price ω_p determines the prior about price . The consumer's response strategy $\{f(P, C), \omega_c\}$ is then a solution to the following*

optimization problem.

$$\{f(P,C),\omega_c\} = \arg\max_{f'(\cdot,\cdot),\omega_c'} \int_p \int_c U(p,c)f'(p,c)\omega_c'(dc)\omega_p(dp), \qquad (5)$$

subject to

$$\int_c f'(p,c)\omega_c'(dc) = g_K(p) \qquad a.s.\omega_p \qquad (6)$$

$$f'(p,c) \geq 0, \qquad \forall p,c \qquad (7)$$

$$I(P;C) \leq \kappa. \qquad (8)$$

$I(P;C)$ *is mutual information between random variables* P *and* C, *defined as*

$$
\begin{aligned}
I(P;C) &= H(P) - H(P|C) = \\
&= \int_{pc}\int_{pc} f'(p,c)\log\Big(\frac{f'(p,c)}{g_K(p)f'(c)}\Big)\omega_c'(dc)\omega_p(dp). \qquad (9)
\end{aligned}
$$

(6) *requires consistency with prior knowledge and* (7) *states the non-negativity of a probability distribution.* (8) *is the information constraint.*

The condition expectation of consumption given a price now takes the form:

$$E[C|p] = \int cf(c|p)\omega_c(dc). \qquad (10)$$

2.2 Monopolistic Seller

In general, the seller could be optimizing his pricing strategy in a class of mixed strategies by choosing a distribution of prices for each single input cost. However, it is shown in Lemma 5 in Appendix B that only a pure pricing strategy can be optimal.

Let a pricing strategy, $\tilde{p}(\mu)$, be a function determining a unique value of the price given a realized unit input cost. The selected strategy maximizes the expectation of profit Π,

$$\Pi = c \cdot (p - \mu), \qquad (11)$$

where c is the sold amount, i.e. the consumer's consumption.

If the consumer observed the price exactly, her consumption would be a function of the realized price only - independent of other prices in the seller's pricing strategy. However, when a consumer is constrained by limited information capacity, consumption also depends on her prior knowledge about price, i.e.

on the whole probability distribution of prices. It was discussed in the previous section that the distribution of prices influences the choice of the consumer's decision strategy. Prior knowledge affects what forms of information the consumer processes and thus partially shapes her final decision on how much to consume. Therefore, the seller does not set optimal prices on a cost by cost basis. The pricing strategy at one cost influences the consumer's behavior at all prices. The seller selects the whole pricing strategy simultaneously, considering the cross-effects of charging a specific price at one given realized input cost on the consumer's responses to other prices charged at different input costs. The overall distribution of prices is determined by the pricing strategy $\tilde{p}$ together with the distribution of unit input costs given by a pdf $h(\mu)$. The probability that the price belongs to a certain interval equals the total probability of input costs that generate prices falling into such an interval. The probability measure of a set A of prices thus equals the following:

$$\omega_p(A) = \int_S h(\mu)d\mu, \qquad S = \{\mu : \tilde{p}(\mu) \in A\}. \tag{12}$$

Definition 2 *__The seller's problem:__ The seller chooses a pricing strategy $\tilde{p}(\mu)$ that maximizes the expected profit.*

$$\tilde{p}(\mu) = \arg\max_{\tilde{p}'} E_{\tilde{p}'}[\Pi] = \arg\max_{\tilde{p}'} \int E[C|\tilde{p}'(\mu)] \cdot (\tilde{p}'(\mu) - \mu)h(\mu)d\mu. \tag{13}$$

$h(\mu)$ is a given distribution of unit input costs. The expected consumption,$E[C|\tilde{p}'(\mu)]$, is given by (10) for a consumer's decision strategy,f, which is a solution to the consumer's problem, (5)-(8), such that prior knowledge coincides with the true distribution of prices, (4), and the distribution of prices is determined by the pricing strategy, (12).

3 Solutions

3.1 Consumer's Response Strategy

This section studies the properties of the solutions to the consumer's problem, $(5) - (8)$. It also discusses how the consumer's responses change when a price distribution becomes more concentrated or when precision of signals on price improves.

If the information constraint (8) is not binding, then the setup takes the form of a standard perfect information maximization problem. Posterior distributions are arbitrarily precise; and they degenerate to delta functions at the true values of price. Conditional distributions of C given p degenerate at $c = c_{opt}(p)$, where $c_{opt}(p)$ is the perfect information demand function.

$$c_{opt}(p) \quad = \quad \arg\max_{c'} U(p, c'), \tag{14}$$

$$\tag{15}$$

For a CES utility, U_1, the perfect-information demand function is

$$c_{opt}(p) = \frac{1}{(1/a - 1)^\theta p^\theta + a^2 p} \tag{16}$$

while for U_2, it is

$$c_{opt}(p) = \left(\frac{\theta}{\theta - 1}\right)^{-\theta} p^{-\theta}. \tag{17}$$

However, it is much more difficult to solve the consumer's problem when constraint (8) binds. I was not able to find analytical solutions to the consumer's problem $(5) - (8)$ in general. Therefore, I rely on semi-quantitative insights and on numerical solutions[5]. The properties of the consumer's responses are summarized in points $C1 - C4$.

C1: The consumer targets $c_{opt}(p)$, but imperfectly. Both perfect information demand functions, (16) and (17), are strictly decreasing and convex. The expected consumption is decreasing even if the consumer has a limited information capacity.

---

[5] The original problem $(5) - (8)$ was discretized into the corresponding finite dimensional convex problem. Details of the numerical method are described in Appendix C
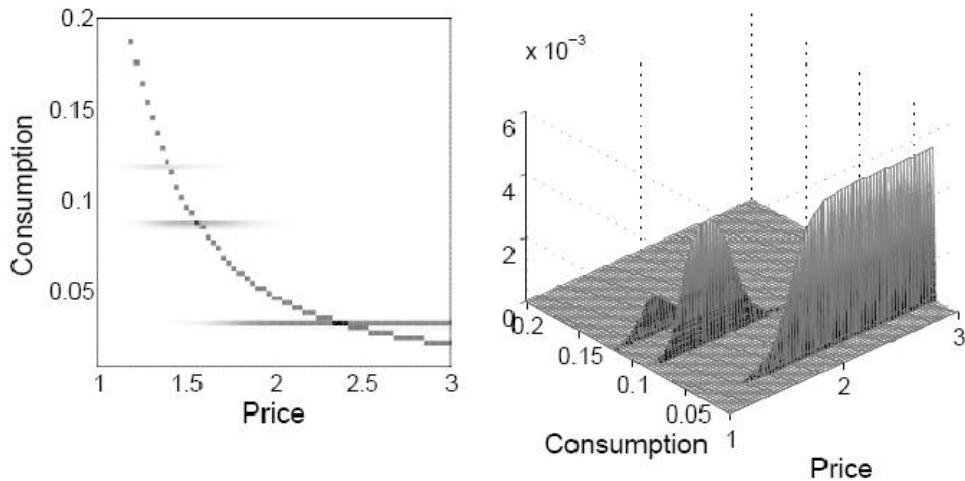
Figure 1: Joint distribution of prices and consumption, $\theta = 3$, $\kappa = 1$ bit

Proposition 1 *If an indirect utility function $U(p,c)$ satisfies $\frac{\partial^2 U}{\partial c \partial p} < 0$ and $\kappa > 0$, then $E[C|p]$ is a strictly decreasing function of p in the support of the prior on p. Both U_1 or U_2 satisfy the assumption above.*

Proof: This is a trivial application of Lemmata 1, 2 and 4 in Appendix A.

Figure 1 shows a numerical solution for the utility function U_2, price uniformly distributed in $(1,3)$, and the information capacity $\kappa = 1$. Both graphs present the same joint pdf $f(p,c)$. The dashed curve in the left graph represents the optimal demand function, $c_{opt}(p)$. On the right, it is well visible that the consumer decides to process three different realizations of signals only. The three signals lead to three different values of consumption[6]. Prices between 1 and 2 usually generate a signal implying the lowest consumption, sometimes they lead to the middle signal; and almost never to the last signal. The last signal carries information that the price is quite likely very high.

The consumer would like to collapse the joint distribution of price and consumption on the $c_{opt}(p)$ curve, but she can not, since she is not able to acquire perfect signals. Imperfect and noisy signals generate sub-optimal responses. The joint distribution is therefore dispersed around the optimal curve.

The higher the consumer's information capacity is, the closer to the optimal

---

[6]Discreteness is discussed in Sims [12], Matějka and Sims [?] and Matějka([6]
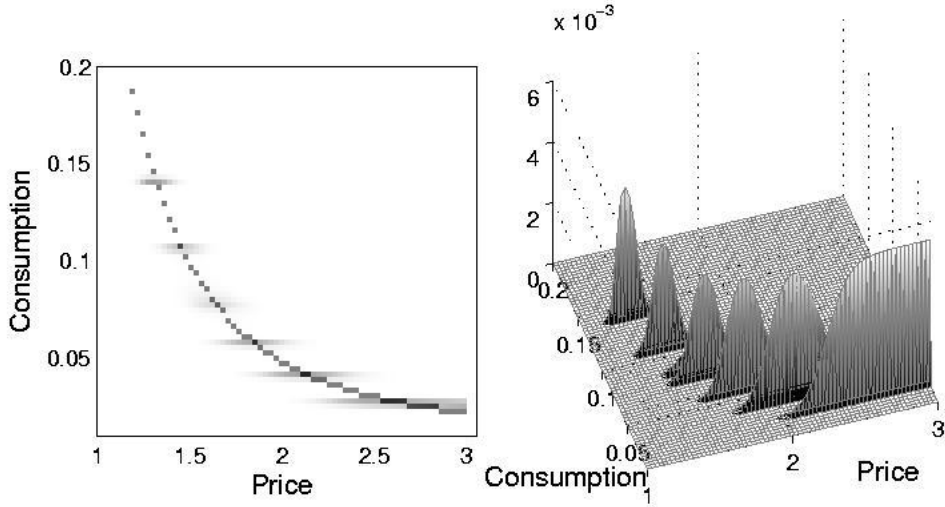
Figure 2: Joint distribution of prices and consumption, $\theta = 3$, $\kappa = 2$ bit

curve her responses are. Figure 2 presents a numerical solution when $\kappa = 2$. The responses may seem quite imperfect. However, in comparison with the solution under perfect information, the consumer loses 1.15% if $\kappa = 1$ and 0.06% when $\kappa = 2$, which corresponds to only 0.6% and 0.03% reductions in consumption[7].

C2: The consumer processes more information at lower values of prices. Figures 1 and 2 show that the consumer chooses to acquire relatively tighter signals when the price is low. She does so to minimize losses from imperfect knowledge; the losses are potential higher at lower prices. It is shown in [6] that an agent processes more information when the loss factor, $L(p)$, is higher.

$$L(p) = -\left(\frac{dc_{opt}(p')}{dp'}\right)^2 \frac{d^2U(p,c)}{dc^2}, \qquad \text{at } c = c_{opt}(p), p' = p. \qquad (18)$$

The change in utility due to misjudging a price p by a small amount ϵ is approximately equal to $-L(p)\epsilon^2/2$. The loss factor takes an especially simple form for the utility function U_2:

$$L(p) \propto p^{-\theta-1}, \qquad (19)$$

---

[7]This amount does not include utility effects of the indirect part $(-pc)$

11

which is a decreasing function of p. Processing information about low prices is more valuable.

C3: Under certain conditions, the consumer chooses to consume less if prices are more dispersed. Lemma 3 in the appendix states that if

$$\frac{\partial U(p,c)}{\partial c} \text{ exists and is concave in } p, \tag{20}$$

then the consumer chooses to consume more when she has acquired a prefect signal on price than when the acquired signal has the same expectation but is more dispersed. In other words, a completely inattentive agent chooses to consume more when price is totally rigid than when it is volatile about the same expectation.

The consumer fears that she loses disproportionately on consuming too little of the other goods, in case the seller's price turns out to be unexpectedly high. Therefore, she rather consumes more of the basket of other goods with a known price.

With a fixed information capacity, or a convex cost of processing information, more dispersed prices and thus prior knowledge imply less accurate posterior knowledge about the price. Therefore, volatile prices drive consumption down. Consumers prefer and reward stable prices.

This effect is an analog to precautionary saving in a standard savings problem with a stochastic stream of endowments. If uncertainty about future endowments increases, precautionary saving increases, too. In our model, the role of future consumption amounts is played by a consumption of the basket - uncertainty about the amount is driven by uncertainty about the seller's price, i.e. by what is going to be left of the initial endowment after purchasing the seller's product.

The assumptions of Lemma 3 are satisfied by the CES utility function,(1); it is verified in the proof of Lemma 4. However, U_2 does not satisfy (20), since $\partial U_2(p,c)/\partial c$ is linear in p.

$$c \;=\; \arg\max_{c'} E_p[U_2(P,c')] = \arg\max_{c'} E_p[c'^{1-1/\theta} - Pc'] \tag{21}$$

$$\Rightarrow c \;=\; (1-1/\theta)^{\theta} E_p[P]^{-\theta}, \tag{22}$$

where $E_p[\cdot]$ is the expectation operator given a distribution of p. For U_2, consumption depends on the expectation $E_p[P]$ only; shape of the distribution has no extra effect on the choice of optimal consumption. U_2 is a scaled version of a limiting case of CES if $a \to 0$; it is when a portion of the initial endowment spent on the seller's product is negligible.

What is really needed for a dispersion of prices to have a negative effect on expected consumption is risk aversion in the spent amount. Let the utility function have, for the sake of simplicity, the following form,

$$U(p,c) = c^r - (pc)^2 \qquad r \in (0,1). \qquad (23)$$

This function differs from U_2 in its second term. Considering utility to be derived from consuming two different goods, the second term represents an ad hoc form of disutility from consuming less of a second good when more of the first is purchased. For U_1, the disutility is linear in the amount spent on the first good if and only if $a = 0$ or $a = 1$.

The optimal amount of consumption for a completely inattentive agent has to satisfy the first order condition:

$$\frac{dE_p[U(P,c)]}{dc} = rc^{r-2} - 2E_p[P^2] = 0. \qquad (24)$$

If no information is processed, the agent's posterior equals her prior, $g(p)$.

$$c = \frac{2}{r}E_p[P^2]^{1/(r-2)} = \frac{2}{r}\left(E_p[P]^2 + Var_p[P]\right)^{1/(r-2)}. \qquad (25)$$

With a fixed expectation of price, consumption is a decreasing function of price's variance.

C4: The consumer's expected consumption increases if she processes more information, especially about lower prices. While the observation $C3$ is concerned with different price distributions, the point $C4$ discusses different collections of signals on the same distribution. $C3$ states that if a seller keeps the expected price fixed and narrows the distribution down, then a consumer will tend to consume more. On the other hand, $C4$ claims that given the same distribution, more is consumed when the consumer acquires tighter signals.

Let a distribution of prices be given. If a demand curve is a strictly convex function of expected price only, then expectation of consumption responses to perfect signals is higher than a consumption response to a mixed signal of all prices. This is a trivial application of Jensen inequality.

The statement can be further refined to any collection of imperfect signals. Expectation of consumption is an integral over all possible signals that can be realized. For the utility function U_2, we get the following:

$$E_s[C] = (1 - 1/\theta)^\theta E_s\left[E_{p|s}[P|s]^{-\theta}\right]. \tag{26}$$

$E_s[\cdot]$ is the expectation operator given a distribution of signals s, while $E_{p|s}[\cdot]$ denotes expectation given a distribution of prices conditioned on a specific signal s. If signals are not perfect, then a strict Jensen inequality holds:

$$\begin{aligned}
(1 - 1/\theta)^\theta E_s\left[E_{p|s}[P|s]^{-\theta}\right] \quad &< \quad (1 - 1/\theta)^\theta E_s\left[E_{p|s}[P^{-\theta}|s]\right] = \\
&= \quad (1 - 1/\theta)^\theta E_p[P^{-\theta}] = E_p[c_{opt}(P)]. \quad (27)
\end{aligned}$$

Therefore,

$$E_s[C] < E_p[c_{opt}(P)]. \tag{28}$$

Responses to perfect signals lead to a higher expected consumption. Although we have shown it analytically for U_2 only, the result is likely to hold for U_1 too, since its main driver is the convexity of a demand function. The higher the convexity, the more extra consumption is generated by additional information processed. As will be discussed later, a seller can respond to this feature of the consumer's behavior by making prices rigid in areas of low convexity of demand (high prices) to save the consumer's information capacity for regions of high convexity (low prices).

$C3$ motivates the seller to keep the price distribution more concentrated. On the other hand, $C4$ motivates him to make high prices more rigid than low ones. He does so to make the consumer especially attentive to price discounts, i.e. to sales.

3.2 Seller's Pricing Strategies

While selecting an optimal pricing strategy, a seller considers the stochastic properties of unit input costs and the consumer's responses to different strategies.

By modifying the pricing strategy together with the whole distribution of prices, the seller stimulates a consumer to process different pieces of information, attain different posterior knowledge and thus respond differently, potentially even to the same realized and imperfectly observed true price.

Let us first inspect optimal pricing in two extreme cases, $\kappa = \infty$ and $\kappa = 0$, which can be studied analytically.

Consumer's information capacity is unlimited, $\kappa = \infty$: The consumer knows the realized price exactly. Her demand for the good depends on its actual price only, not on the whole distribution of prices. Demand always equals the consumer's optimal demand, $c_{opt}(p)$. Therefore, the seller sets an optimal price for each cost separately, he does not need to consider the implications of the shape of the prior distribution for the shape of the posterior.

For instance, if $\theta = 2$ and $a = 1/2$, given an input cost μ, then the maximal profit is achieved by $p = \mu + \sqrt{\mu + \mu^2}$. For $a \in (0, 1)$, the optimal pricing strategy is

$$\tilde{p}(\mu) = \mu + \sqrt{\left(\frac{a}{1-a}\right)^2 \mu + \mu^2},$$ (29)

while the optimal pricing for the utility function U_2 is

$$\tilde{p}(\mu) = \frac{\theta}{\theta - 1}\mu.$$ (30)

Both strategies, (29) and (30), are continuous and strictly increasing functions for all non-negative costs.

Zero information capacity, $\kappa = 0$: The consumer can not process any information. Her posterior knowledge equals her prior knowledge given by the overall distribution of prices. Therefore, given a prior, the posterior knowledge is always the same, regardless of what the actual price is - the seller fully deter-

mines the consumer's posterior knowledge. Optimal prices for different input costs cannot be set independently of each other.

The following proposition is an implication of Lemma 3 in the appendix.

Proposition 2 *Let a utility function be CES, U_1. If $\kappa = 0$, then an optimal $\tilde{p}$ must be constant - the seller charges the same price for all unit input costs.*

Proof: Let us assume that a non-constant $\tilde{p}_1$ is an optimal strategy. Such a strategy delivers a non-degenerate distribution of prices, with a pdf $g_1(p)$. Since the consumer has zero information capacity, her posterior knowledge equals her prior (its pdf also equals $g_1(p)$). Given this posterior knowledge, she chooses an amount c to maximize expected utility. With no information specific to the actual price p, the consumer always selects the same c, regardless of the realized price.

Let $\tilde{p}_2(\mu) = \bar{p}$ be an alternative pricing strategy, where $\bar{p}$ is the mean price of $\tilde{p}_1$. Let c' be a consumption that is realized when $\tilde{p}_2$ is applied. Lemma 3 in the appendix states that

$$c < c'. \tag{31}$$

Consumption rises if the constant pricing strategy is used. Moreover,

$$
\begin{aligned}
E_{\tilde{p}_1}[\Pi] &= c \cdot E_{\tilde{p}_1}[p - \mu] < \\
&< c' \cdot E_{\tilde{p}_1}[p - \mu] = c' \cdot (\bar{p} - \bar{\mu}) = E_{\tilde{p}_2}[\Pi], \tag{32}
\end{aligned}
$$

so the seller's expected profit rises, too. No non-constant pricing can be optimal.QED.

These two extreme cases illustrate two main forces acting on the choice of the pricing strategy. If the information constraint is binding, the consumer's response to each single cost depends on the whole distribution of prices, too. The seller then tends to choose more condensed pricing, realizing that consumption is likely to fall when the signal on price is more dispersed. This force reflects the property $C3$ - the consumer's cautiousness.

On the other hand, any time the information capacity is positive, the consumer does acquire some refined knowledge about the actual price. The second force realizes the consumer's demand as a function of the level of the price - it

makes the seller desire to price differently for different input costs. This force dominates at higher information capacities, and is the only one at play when $\kappa = \infty$, while cautiousness is the main driver if $\kappa = 0$, when the consumer can not distinguish between prices at all.

Setups with a finite information capacity were studied numerically. Analytical solutions are not in general available even just for the consumer's problem under constraints on information capacity. So, it is well out of reach at the moment to study the seller's optimization over pricing strategies when coupled with the problem of an inattentive consumer. Numerical methods that were applied are described in the appendix.

Finite information capacity, CES utility. Figure 3 shows numerical solutions for CES with $\theta = 2$ and four different levels of information capacity. The solid line in each graphs represents a perfect-information pricing strategy, (29). The optimal pricing strategies are imperfectly concentrated around this line. The higher the information capacity, the closer the seller's strategy and the perfect-information strategy are.

The upper left picture, when a consumer is completely inattentive, is a manifestation of Proposition 2 - the optimal pricing strategy is constant and the price is completely rigid. If the information capacity increases, pricing becomes more flexible. For $\kappa = 1$, the seller chooses to charge two different values, for $\kappa = 2$ it is four values. Pricing is completely flexible when the capacity is unlimited; a numerical solution tracks the analytical solution (29) very closely.

In all four cases, the seller chooses a strategy of maximal entropy such that a consumer can still observe price exactly. 1 bit of information capacity allows the consumer to distinguish between two different values, while 2 bits distinguish between four of them. The seller recognizes that a more concentrated prior allows tighter posterior knowledge, which in the case of CES leads to higher consumption. In general, we should expect that the more attentive the consumer is, the more flexible pricing strategy the seller selects.

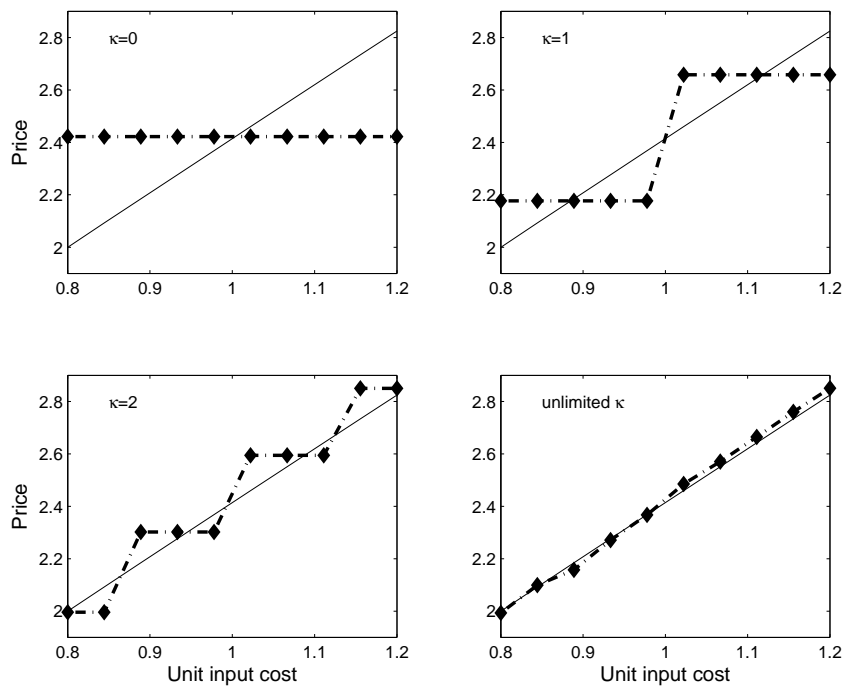*Finite information capacity, utility function U_2.* Optimal pricing strategies
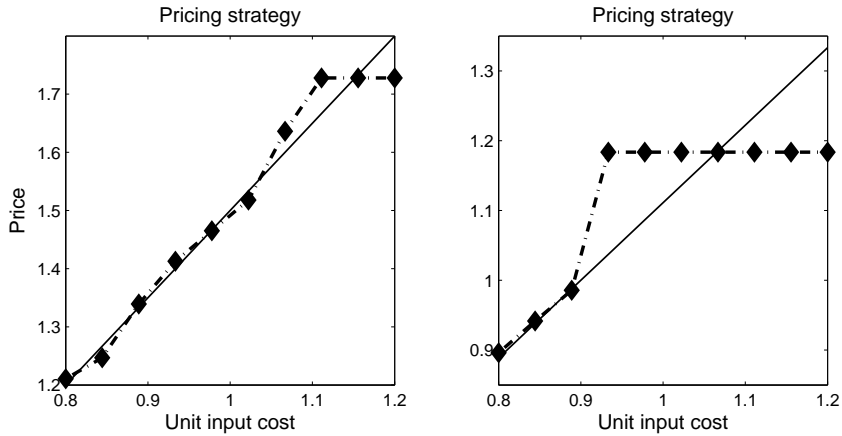
Figure 3: Pricing strategies, CES, $\theta = 2$

Figure 4: Pricing strategies, U_2, $\kappa = 1$, left: $\theta = 3$, right: $\theta = 10$

for the utility function U_2,(2), are shown in Figure 4. Information capacity is 1 bit. Pricing is less rigid than for CES in Figure 3, but it is still not completely flexible. U_2 does not satisfy the assumptions of Lemma 3. Making the whole prior more concentrated does not generate higher consumption. Unlike for CES, pricing is not rigid at all input costs.

However, the seller does prefer to price rigidly at the highest input costs. Selecting one high price instead of several of them prevents a consumer from processing information about which one of the high prices is the realized one. The seller does so to stimulate a consumer to use most of her information capacity at lower prices. $C4$ discusses that consumption increases if signals become tighter. The effect is stronger if the convexity of a demand function is higher. This feature of the consumer's behavior motivates the seller to vary the flexibility of his strategy between different levels of input costs. Rigid pricing in regions of low demand convexity saves on the consumer's information capacity, which will then be used in regions of high convexity. For U_2, convexity is a decreasing function of price.

Rigid high prices allow the consumer to use more information capacity when the price is low. In other words, the consumer can pay more attention to sales, which in turn benefits the seller.

Finite information capacity, LQ utility. Let the utility have the following
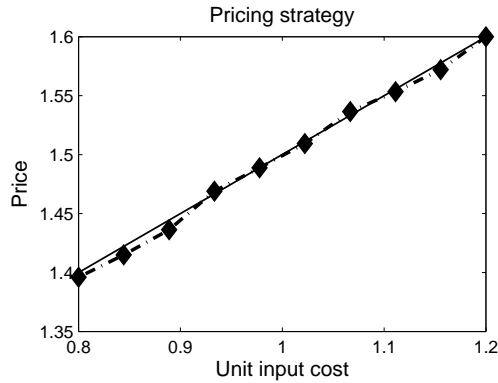
19

Figure 5: Pricing strategies, LQ utility function, $\kappa = 1$

form:

$$U(p, c) = -c^2 + 2c - pc. \tag{33}$$

This utility does not satisfy the assumptions of Proposition 2, moreover, the convexity of the induced demand function is constant,

$$c_{opt}(p) = 1 - \frac{p}{2}. \tag{34}$$

None of the drivers towards pricing rigidity mentioned above are present. Figure 5 presents a numerical solution for $\kappa = 1$ and the analytical solution for $\kappa = \infty$. Pricing is completely flexible even if the consumer's information capacity is low,

$$\tilde{p}(\mu) = 1 + \frac{\mu}{2}. \tag{35}$$

Let us summarize the discussion above.

S1: The seller's pricing strategy is an imperfect approximation of his perfect-information strategy.

S2: Preference for discreteness: there is is an apparent tendency toward rigidity of prices. Only a few distinct values are to be charged even when unit input cost takes values in a continuous interval.

S3: The rigidity of pricing is higher if the consumer's information capacity is lower: rationally inattentive consumers reward stable prices.

20

S4: The rigidity of high prices increases if the convexity of the demand function decreases with prices more rapidly: the importance of proper attention to sales increases; the seller accommodates the consumer to process more information about low prices.

4   Dynamic Models

This section outlines the potential implications of consumers' inattention for dynamic models of pricing. A complete formalization of such models is unfortunately well beyond a scope of this paper. A proper dynamic extension of the static model presented earlier would be too demanding to solve even numerically. While in the static version, the seller's strategy is a function on a space of unit input costs, strategies in the dynamic version would take the form of functionals on a space of all the paths of unit input costs to a space of all the price paths. Optimal prices can turn out to be serially correlated even if the input cost is i.i.d. Instead of constructing a fully-fledged dynamic model, we set up a couple of illustrative models.

1. The following section presents a model with a serially correlated unit input cost yielding rigid prices. The drivers of these results are the same as those presented in the analysis of the static model: consumers reward stable prices with increased consumption. The seller chooses to price discretely to make prices easily observable. Although this model is again based on imperfect processing of information, we make simplifying assumptions that prevent consumers from processing information gradually. Prior knowledge about prices is fixed in time.

2. The next section, on the other hand, sheds light on the implications of the gradual refinement of knowledge in time. There is either a positive or a negative shock to an aggregate component of the seller's input cost in period 0. Consumers have limited information capacity and learn about the shock's true realization only gradually, exogenously of the seller's actions. It turns out that the seller prices in line with the consumers' expectations: he changes his price as sluggishly as consumers learn about the shock. The seller does so to better utilize consumers' knowledge and decrease their uncertainty about the price.

   Drawing on these results, we conclude the section by discussing implications of rational inattention for more complex models.

4.1 Discrete Prices in a Dynamic Setting

This model, DynA, is going to take the same analytic form as the original static model formulated in Definition 2. Assumptions I make will imply that consumers keep their priors about prices fixed, which will make the consumer's side of the problem static.

A consumer is periodically endowed with a nominal endowment, e. She first visits the seller's store and then automatically spends the rest of the endowment on other products during the time that is remaining until the next endowment. There is no savings beyond the moment of the next endowment. The number of consumers arriving to the store each period is constant. The time intervals between successive endowments have the same length for all consumers, but the phases of their arrivals are not synchronized.

I make the following assumptions.

i) Consumers process information about the seller's price when they visit the store only.

ii) Unit input cost changes on a much quicker time scale than that on which consumers visit the store.

iii) The stochastic process of unit input costs is ergodic. The limiting distribution of the unit input cost has bounded support and is independent of current and past unit input costs. For instance, the cost does not have a long-term trend. However, the input cost can be i.i.d. or serially correlated.

The seller incurs a stochastic unit input cost and applies his pre-selected pricing strategy. The strategy determining the price to be charged could in general be a function of current as well as past unit input costs. However, due to the assumptions above, the consumer's prior knowledge about the price at the moment when she enters the store is independent of any information she processed in the past - it always coincides with the limiting (long-term) distribution of prices, which is fixed. In such a case, the seller's optimal pricing strategy is a function of the current unit input cost only. The strategy together with the limiting distribution of costs determines the limiting distribution of prices.
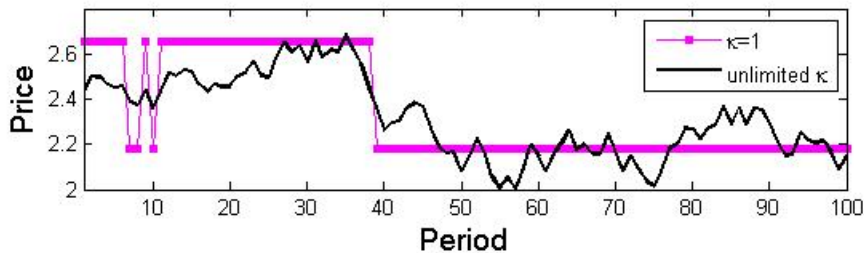
Figure 6: Simulated price series, CES, $\theta = 2$, $a = 0.5$, stable cost

Although the setting was originally dynamic, the consumer's problem again takes the form of (5)-(8), where prior g is the limiting distribution of prices. The seller chooses an optimal $\tilde{p}(\mu)$. Realized prices are then determined by successive application of $\tilde{p}(\mu)$, even if the unit input cost, μ, is serially correlated.

Definition 3 *Model DynA: takes the same form as the seller's problem formulated in Definition 2, except that $h(\mu)$ is the limiting distribution of input costs:*

$$h(\mu) = \lim_{T \to \infty} \frac{1}{T} = \int_0^T h(\mu, t)dt, \tag{36}$$

where $h(\mu, t)$ is a prior on costs at time t, which is determined by the properties of the stochastic process of the unit input cost.

A simulated price series for $\theta = 2$ and $a = 0.5$ is shown in Figure 6. The picture compares optimal pricing when consumer's information capacity is $\kappa = 1$ with fully flexible pricing at $\kappa = \infty$. The series of flexible prices fully reveals the realized unit input costs, while optimal prices when $\kappa = 1$ take two different values only. The lower value of the price is selected anytime μ drops below 1.

I simply generated a serially correlated μ from a stochastic process with a long-term distribution that is uniform in $(0.8, 1.2)$ [8], and then successively applied the appropriate pricing strategy, which is presented in the upper right graph of Figure 3.

Figure 8 shows a simulated time series of prices for $a = 0.25$ - the corresponding optimal pricing strategy is on the left in Figure 7. A lower share parameter implies more flexible pricing; prices take 3 different values. Finally,

---

[8]For purposes of the simulation, it is only the specific form of the cost-path realization that is important, not the properties of the underlining stochastic process.
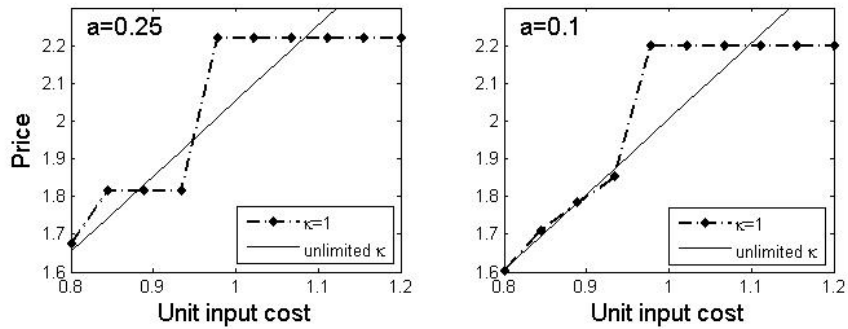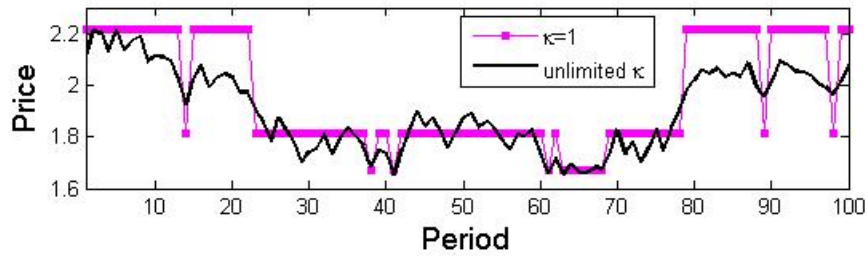
Figure 7: Pricing strategies, CES, $\theta = 2$

Figure 8: Simulated price series, CES, $\theta = 2$, $a = 0.25$, stable cost

Figure 9 presents price series for $a = 0.25$, but this time the underlining unit input cost is more volatile. The resulting prices change their values more often.

4.2   Delayed and Smoothed Adjustments to Aggregate Shocks

Rationally inattentive agents learn about new innovations slowly. While the gradual refinement of information was completely neglected in the previous sec-
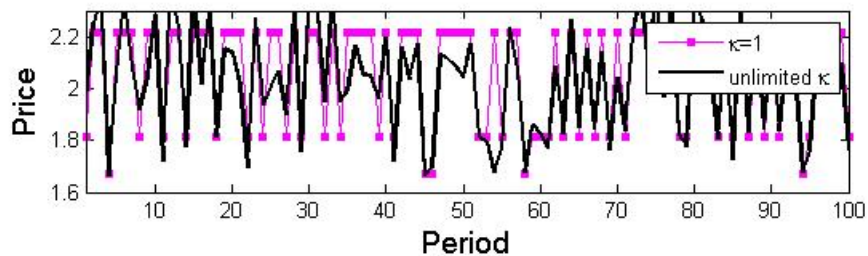

Figure 9: Simulated price series, CES, $\theta = 2$, $a = 0.25$, volatile cost

25

tion (consumers did not learn at all), it is the main theme in this one. Consumers gradually refine their knowledge about stochastic variables. If there is a shock to an aggregate variable and if the seller's input cost is correlated with this variable, then the seller chooses to respond to such a shock gradually - he chooses to price in line with the consumer's expectations.

Imagine a consumer has some partial knowledge about shocks to energy prices. She also knows that energy prices are the main determinant of the input cost for her favorite local sauna club. The consumer's expectations about the admission prices to the sauna vary with what she knows about the current prices of energy. The sauna owner might postpone new price changes until consumers expect them to occur.

The model DynB has these features:

i) The input cost is drawn from a binary distribution in the period 0.

ii) The consumer's knowledge of the seller's cost evolves independently of the seller's actions. Knowledge is gradually refined.

iii) The seller's price is a function of the unit input cost and the time elapsed from the initial shock.

Let the seller's unit input cost be equal to an aggregate variable A. The seller is small and has a negligible influence on the consumer's knowledge about shocks to A - the knowledge evolves independently from the seller's pricing responses to A. A is drawn from a symmetric binary distribution $\{A_L, A_H\}$ in the period 0 and stays constant forever after. Consumers know that one of the two possible shocks is realized, but need to process information to find out which one it is. Such a setting with a one-time shock is both simpler to solve and yet illustrative enough to document the implications of gradual knowledge adjustment.

In fully-fledged models under rational inattention, we would specify consumers' preferences and allow them to choose what pieces of information to process. I will, however, assume one specific form of information structure. The qualitative properties of the results do not rest on this assumption.

Let us assume that the consumer's knowledge in period t has the same form as if the consumer acquired one signal through a binary channel with a noise

level $X(t)$. $X(t)$ is decreasing in t, which models knowledge refinement[9]. With increasing time, there is a higher probability that agents receive the correct signal. Posterior knowledge is thus more concentrated.

If $A = A_H$, then the probability that an agent receives the correct signal, $(A = A_H)$, is $1 - X(t)$. The posterior knowledge of an agent having received such a signal is $\{\mathbb{P}(A_L) = X(t), \mathbb{P}(A_H) = 1 - X(t)\}$. The posterior knowledge of agents who received the corrupted signal, $(A = A_L)$, is $\{\mathbb{P}(A_L) = 1 - X(t), \mathbb{P}(A_H) = X(t)\}$.

The seller chooses his pricing strategy. Unlike in the earlier sections of this paper, the strategy is not a function of the input cost only. The consumer's knowledge evolves even after period 0, when the input cost is kept fixed. Different consumer's knowledge can imply a different optimal pricing response to the same input cost. The pricing strategy takes the form $p = \tilde{p}(A, \{g(A)\}) = \tilde{p}(A, t)$, where $\{g(A)\}$ is the distribution of knowledge[10] in the population of consumers. $\{g(A)\}$ is determined by $X(t)$, which is pinned down by time t. The strategy can be expressed using two functions, $\tilde{p}_L(t)$ and $\tilde{p}_H(t)$, each corresponding to one level of unit input cost:

$$
p = \begin{cases} \tilde{p}_L(t) & \text{if } A = A_L, \\ \tilde{p}_H(t) & \text{if } A = A_H. \end{cases} \tag{37}
$$

Consumers are rational, they know the form of $\tilde{p}_L(t)$ and $\tilde{p}_H(t)$. Together with their knowledge about the aggregate shock (A determines which one of $\tilde{p}_L(t)$ and $\tilde{p}_L(t)$ is to be applied), the pricing strategy generates the consumer's prior on price. More specifically, the pricing strategy forms the prior's support, while knowledge about A determines the relative probabilities of its two points. The term prior reflects knowledge "before" a consumer processes information about the seller's price, but it is "after" she has processed information about the aggregate shock. Since the consumer's knowledge about A is independent of the seller's actions, optimal prices at different levels of $X(t)$ can be set independently of each other.

---

[9]No sequence of signals across periods is considered, just one signal, which gets tighter in latter periods.

[10]It is actually a distribution of distributions.

Definition 4 *Model DynB:* $X(t)$ *is given for all* $t \in \{0..\infty\}$; *it is a non-increasing function . For each* t, *the seller chooses* $\tilde{p}_L(t)$ *and* $\tilde{p}_H(t)$, *maximizing the expectation of his profit*

$$
\begin{aligned}
\{\tilde{p}_L(t), \tilde{p}_H(t)\} \quad &= \quad \arg\max_{\{\tilde{p}'_L(t), \tilde{p}'_H(t)\}} \\
&\Big\{ \quad \tfrac{1}{2} \quad (\tilde{p}'_L(t) - A_L) \\
&\qquad\qquad \Big((1 - X(t)) E[C|t, A_L] + X(t) E[C|t, A_H] \Big) \\
&\quad + \quad \tfrac{1}{2} \quad (\tilde{p}'_H(t) - A_H) \\
&\qquad\qquad \Big(X(t) E[C|t, A_L] + (1 - X(t)) E[C|t, A_H] \Big) \Big\}. \quad (38)
\end{aligned}
$$

The expression for expected profit weights the true realizations of A_L *and* A_H *and also the consumer's priors generated by receiving signals on* A_L *or* A_H. $E[C|t, A_L]$ *denotes the consumption expectation when a consumer's prior is determined by a signal pointing to* A_L *with a noise level* $X(t)$ *- the corresponding prior is* $\{g(\tilde{p}_L(t)), g(\tilde{p}_L(t))\} = \{1 - X(t), X(t)\}$. *On the other hand, a prior determining* $E[C|t, A_H]$ *is* $\{1 - X(t), X(t)\}$. *The consumption expectation is evaluated from a solution to the original consumer's problem (5)-(8) with the appropriate prior.*

For each $X(t)$, the numerical representation of the optimal pricing strategy can be found simply by evaluating the expected profit for all combinations of $\{\tilde{p}_L(t), \tilde{p}_H(t)\}$. Let noise decrease at the following rate:

$$
X(t) = 0.5 - 0.05t, \qquad \forall t \in \{0..10\}. \tag{39}
$$

If the realized value is $A = A_H$, then the seller's price gradually increases until it reaches the full information price in period 10. Otherwise, it gradually decreases. For simplicity, let $\kappa = 0$. Consumers do not process any additional information about the seller's price, they only use their knowledge about the aggregate variable. If $\kappa > 0$, the same optimal prices would correspond to higher levels of X.

Figure 10 presents the optimal pricing strategies as a function of time t, for CES utility(U_1), $a = 0.5$, $\theta = 2$ $A_L = 0.8$ and $A_H = 1.2$.

Consumers possess very little knowledge about A in early periods. They know the seller's pricing strategy, but have difficulties distinguishing between
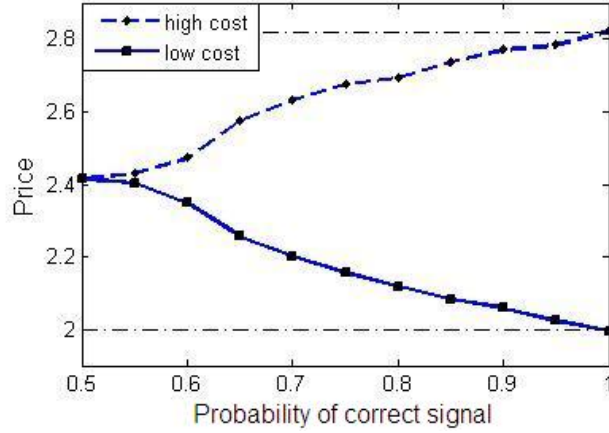
Figure 10: Gradual price adjustment to aggregate shock, $\theta = 2$.

the two different values of prices, $\tilde{p}_L(t)$ and $\tilde{p}_H(t)$, that can be realized in the particular period. If $X(t) = 0$, consumers always acquire the correct signal, then the seller sets the perfect information optimal prices, which are represented by the dash-dotted bounds. If $X(t) = 0.5$, consumers can not tell at all which of the two prices was realized - in such a case, the seller chooses to set one price only. Like in the static model, consumers consume more when they are less uncertain about prices. With the increasing probability of the correct signal, optimal prices $\tilde{p}_L(t)$ and $\tilde{p}_H(t)$ are set further and further away from each other.

Figure 11 shows the same solution with the x-axes scaled differently. In this case, it is the amount of information capacity used that increases linearly with time rather than the level of noise decreasing linearly. This scale is probably the more natural one.

$$\kappa(t) = t/10, \qquad \forall t \in \{0..10\}, \tag{40}$$

$$\kappa(t) = 1 + X(t)\log X(t) + (1 - X(t))\log(1 - X(t)). \tag{41}$$

The second equation relates the capacity of a binary channel to the resulting level of noise. The noise level, $X(t)$, is a convex function of information capacity. Very little capacity is needed to decrease the noise from 0.5 to 0.45, a bit more from 0.45 to 0.4, etc.
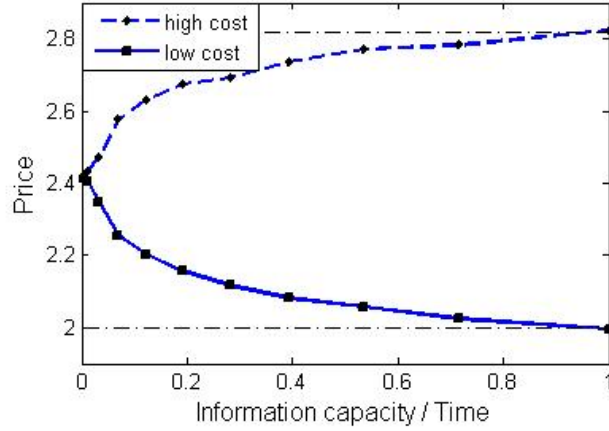
Figure 11: Gradual price adjustment, scaled to information amount

4.3 Implications

4.3.1 Staggered Price Changes

Under rational inattention, price changes can also be sluggish and yet staggered, unlike in DynB, where price responses were smooth.

A support $\mathcal{S}$ of the consumer's prior consists of price-points p that lie in the range of the pricing function $\tilde{p}(\mu)$.

$$\mathcal{S} = \Big\{ p : \exists \mu; \tilde{p}(\mu) = p \Big\}. \tag{42}$$

The seller desires to limit the prior's entropy. In DynA, he chooses to charge the same price for intervals of the unit input cost.

In DynB, priors evolve according to the evolving pricing strategy $\tilde{p}(t, \mu)$. The time component is present due to consumers' gradual collection of information about the period-zero shock. Consumers know the seller's pricing strategy as a function of time and also how much time has elapsed since the moment the shock occurred. Therefore, they construct their priors directly from price points that can be charged at that specific period.

$$\mathcal{S}(t) = \Big\{ p : \exists \mu; \tilde{p}(t, \mu) = p \Big\}. \tag{43}$$

Even if DynB was solved for a continuous interval of unit input costs, $\tilde{p}(t, \mu)$ would likely be a piece-wise constant function of μ (at a fixed time t). However,

30

since $\tilde{p}$ can vary continuously with time t, the impulse responses of realized prices can be smooth, just like in Figure 10.

Optimal pricing strategies in more complex models with rationally inattentive consumers would, however, likely feature staggered rather than smooth price adjustments. This can arise if consumers can not fully identify the exact moment when a shock occurred. For instance, we could model an infinite horizon problem with potential shocks in each period and assume that consumers do gradually learn about shocks just like in DynB, but they can misjudge the moment of a shock by at most one period. If successive shocks are always distant enough, so they do not interfere, then the sellers pick a pricing strategy as a function of the shock's form and of a time elapsed from the shock. In this case, the consumer's prior takes the following form.

$$\mathcal{S}(t) = \left\{ p : \exists \mu, \exists \Delta \in \{-1, 0, 1\}; \tilde{p}(t + \Delta, \mu) = p \right\}, \tag{44}$$

where $t-1, t, t+1$ are possible numbers of periods from the shock as perceived by the consumer. The support is generated by price values corresponding to three different time levels. In order to decrease the support's entropy, the seller needs to coordinate price values across periods. Moreover, price points in all periods are bound together through the succession of these three-period kernels. A possible solution is a pricing strategy gradually responding to new shocks yet constant on finite time intervals. Price responses would be delayed and sluggish, but spread over a couple of discrete jumps only.

What we really need to get the staggered price adjustment is a noise component in the rate of information acquisition. However, this component is present if agents process information through channels of limited information capacity. It is only the average rate of information that is bounded by the agent's information capacity, but realized rates do deviate from this quantity.

4.3.2 Rigidity of Real and Nominal Variables

The results presented in this section have another appealing implication. This is that consumer's inattentiveness suffices to generate both nominal and real rigidities. As Sims[10] notes, this overall rigidity is an important feature of data that both classical as well as early New Keynesian models fail to generate. Al-

though more recent New Keynesian models do display real rigidity, they usually achieve it via explicit and often ad hoc assumptions of real frictions.

Sims claims that models with inattentive agents on the supply and the demand side of the economy would reconcile with the observed general rigidity of variables. Early New Keynesian models postulating rigidity of prices typically imply an immediate increase in output as a response to monetary expansion. The expansion of monetary aggregates raises nominal demand, which under rigid prices also increases demand. Output is then assumed to satisfy demand exactly, and thus to immediately respond to nominal shocks.

It is straightforward to realize that introducing a constraint on the consumer's information capacity in a New Keynesian model would generate real rigidity. If a consumer chooses a level of consumption while processing information somewhat slowly, then she can not adjust her demand instantaneously after a shock. She responds only as quickly as she learns about the shock.

Moreover, this paper discusses that such a constraint on the consumer's side generates nominal rigidity, too. However, nominal movements are in fact in the seller's discretion. The consumer's inattention is therefore a driver of the rigidity of both nominal as well as real variables.

5 Conclusion

This paper proposes a mechanism leading to rigid pricing as an optimal strategy. It shows that consumers' limited abilities to process information can motivate sellers to keep their prices rigid.

The presented models apply the framework of rational inattention to study the pricing strategies of a monopolistic seller facing a consumer with limited information capacity. The consumer needs to process information about prices, while the seller is perfectly attentive. It turns out that the seller chooses to price discretely even for a continuous range of unit input costs, i.e. the seller charges a finite set of different prices only. He does so to provide the consumer with easily observable prices and thus stimulate her to consume more. In dynamic models, this mechanism, in addition, implies that prices respond to cost shocks sluggishly.

Solutions to the model's dynamic version, the price series presented in Figures 6 and 8, resemble some real time series of prices quite well[11]. Prices often stay constant for a while and attain just a few different values in the long run. The frequency of switching between different values can be both high or low depending on the volatility of the underlining unit input cost.

The model's dynamic version generates sluggish responses of prices to aggregate shocks. Sellers wait with a price change until consumers learn about the nature of the shock in order not to surprise consumers with the change. It is also implied that consumers' inattention can generate rigidity of both nominal as well as real variables. This feature is in agreement with the data.

Limited information capacity has been a driver of rigidities in a number of recent papers. However, in this one, it is the consumer who finds it difficult to process information, not the seller. This interaction between a price-setting seller and a rationally inattentive consumer makes the model possible to solve in very simple settings only. On the other hand, the information constraint on the consumer's, rather than the seller's, side certainly has some intuitive appeal - especially in markets where sellers offer just one product, but consumers need to compare the prices of several different sellers.

---

[11]See for example Eichenbaum, Jaimovich and Rebelo [2].

Appendices

A Consumer

Definition 5 *Let $h(c)$ be a marginal distribution of c. A pair of functions $\{F_1(c), F_2(c)\}$ has the m-property over $h(c)$, iff $\exists m; s.t. \forall c, h(c) > 0 : F_1(c) < F_2(c)$ if $c < m$, and $F_1(c) > F_2(c)$ if $c > m$.*

Lemma 1 *Let $f(c|p)$ be a conditional distribution of a solution to (5)-(8) and let $p_1 < p_2$, s.t $g(p_1) > 0$ and $g(p_2) > 0$. If $\frac{\partial^2 U(p,c)}{\partial c \partial p} < 0$, then $\{F_1(c) := f(c|p_1), F_2(c) := f(c|p_2)\}$ have the m-property.*

Proof: A first order condition for a solution of (5)-(8) is

$$f(c|p) = h(c) \exp(U(p,c)/\lambda) w(p), \tag{45}$$

where $h(c)$ is the marginal distribution of c, $\lambda > 0$ is a Lagrange multiplier on the information constraint, (8), and $w(p)$ is a normalization parameter. If $g(p) > 0$, then $w(p) > 0$.

$\frac{\partial^2 U(p,c)}{\partial c \partial p} < 0$ and $p_1 < p_2$ implies that $\Big(U(p_1,c) - U(p_2,c)\Big)$ is increasing in c; it can cross zero at most once. Adding a constant independent of c, we find that $\Big(U(p_1,c)/\lambda + \log(w(p_1))\Big) - \Big(U(p_2,c)/\lambda - \log w(p_2)\Big)$ also has to be increasing and crosses zero at most once. Since exponentiation is a monotone transformation,

$$\exp(U(p_1,c)/\lambda) w(p_1) - \exp(U(p_2,c)/\lambda) w(p_2) \tag{46}$$

also has at most one 0 and is of the opposite sign on either side of some value m for c. But from (45) we know that the above expression, multiplied by $h(c)$, is just $f(c|p_1) - f(c|p_2)$. Therefore over the set of c values at which $h(c) > 0$, $f(c|p_1) - f(c|p_2)$ has the claimed properties. The pair $\{F_1(c) = f(c|p_1), F_2(c) = f(c|p_2)\}$ has the m-property. QED.

Lemma 2 *If a pair $\{F_1(c) = f(c|p_1), F_2(c) = f(c|p_2)\}$ has the m-property and if $\kappa > 0$, then $E[C|p]$ is a decreasing function of p.*

Proof:

$$E[C|p_1] - E[C|p_2] \;=\; \int_{c<m} (f(c|p_1) - f(c|p_2))\, c\omega_c(dc) + \int_{c>m} (f(c|p_1) - f(c|p_2))\, c\omega_c(dc) \geq$$

$$\geq \int_{c<m} \left(f(c|p_1) - f(c|p_2) \right) m\omega_c(dc) + \int_{c>m} \left(f(c|p_2) - f(c|p_1) \right) c\omega_c(dc) =$$

$$= -\int_{c>m} \left(f(c|p_1) - f(c|p_2) \right) m\omega_c(dc) + \int_{c>m} \left(f(c|p_1) - f(c|p_2) \right) c\omega_c(dc) \geq$$

$$= \int_{c>m} \left(f(c|p_1) - f(c|p_2) \right) (c - m)\omega_c(dc) \geq 0. \tag{47}$$

Both inequalities hold assuming the m-property. In the last equality, we used that $f(c|p)$ integrates to 1 for any p. If $\kappa > 0$, then a distribution of c is non-degenerate and the strict inequality holds. QED.

Lemma 3 *If $U(p,c)$ is concave in c, and $\frac{\partial U(p,c)}{\partial c}$ exists and is strictly concave in p, then a perfect signal on a price p^* leads to a higher optimal c then an imperfect signal $h(p)$ with the same expectation equal to p^*.*

Proof: The perfect signal on p^* leads to a consumption amount c^*, which satisfies

$$\left(\frac{\partial U(p,c)}{\partial c} \right)_{c=c^*} = 0. \tag{48}$$

The optimal response c^{**} to the imperfect signal $h(p)$ satisfies

$$E_{h(p)} \left[\frac{\partial U(p,c)}{\partial c} \right]_{c=c^{**}} = 0. \tag{49}$$

Since $\frac{\partial U(p,c)}{\partial c}$ is assumed to be strictly concave, the following holds due to the Jensen inequality

$$0 = E_{h(p)} \left[\frac{\partial U(p,c)}{\partial c} \right]_{c=c^{**}} < \left(\frac{\partial U(p^*,c)}{\partial c} \right)_{c=c^{**}}, \tag{50}$$

which by comparing with (48) implies:

$$\left(\frac{\partial U(p,c)}{\partial c} \right)_{c=c^*} < \left(\frac{\partial U(p,c)}{\partial c} \right)_{c=c^{**}}. \tag{51}$$

The last inequality together with the concavity of $U(p,c)$ in c implies

$$c^* > c^{**}. \tag{52}$$

QED.

Lemma 4 *(a) If $p > 0, c > 0, a \in (0,1), e > cp$ and $\theta \in (1,\infty)$, then the CES utility U_1, (1), satisfies the assumptions of both Lemmata 1 and 3.*

(b) If $p > 0, c > 0$ and $\theta \in (1,\infty)$, then the utility U_2, (2), satisfies the assumptions of Lemma 1 but not of Lemma 3.

Proof: let $r = 1 - 1/\theta$, $r \in (0,1)$. All multiplicative terms in the following expressions are positive.

$$
\begin{aligned}
\frac{\partial^2 U_1(p,c)}{\partial c \partial p} &= -\Big[(1-a)(e-cp)^{r-2}\Big(ac^r + (1-a)(e-cp)^r\Big)^{1/r} \\
&\quad \times \Big((1-a)(e-cp)^{r+1} + c^r a(e-cp) + c^r ae(1-r)\Big)\Big] \\
&\quad \times \frac{1}{\Big(ac^r + (1-a)(e-cp)^r\Big)^2} \\
&< 0,
\end{aligned}
\tag{53}
$$

$$
\begin{aligned}
\frac{\partial^2 U_1(p,c)}{\partial c^2} &= -\frac{(1-a)ac^{r-2}(1-r)e^2(e-pc)^{r-2}\Big(ac^r + (1-a)(e-pc)^r\Big)^{1/r}}{(ac^r + (1-a)(e-pc)^r)^2} \\
&< 0,
\end{aligned}
\tag{54}
$$

$$
\begin{aligned}
\frac{\partial^2}{\partial p^2}\Big(\frac{U_1(p,c)}{\partial c}\Big) &= -(1-a)(1-r)ac^{1+r}(e-pc)^{r-1}\Big(ac^r + (1-a)(e-pc)^r\Big)^{1/r} \\
&\quad \times \frac{\Big((1-a)r(e-pc)^{r+1} + 2(1-a)re(e-pc)^r + ac^r(3e - pc - re)\Big)}{\Big(ac^r + (1-a)(e-pc)^r\Big)^3} \\
&< 0.
\end{aligned}
\tag{55}
$$

$$
\frac{\partial^2 U_2(p,c)}{\partial c \partial p} = -1.
\tag{56}
$$

However, U_2 is linear in p:

$$
\frac{\partial^2}{\partial p^2}\Big(\frac{U_2(p,c)}{\partial c}\Big) = 0.
\tag{57}
$$

QED.

B   Seller

The seller's pricing strategy defines what prices are to be charged at a given realized unit input cost. Such a strategy, potentially mixed, is given by a collection of conditional distributions $H(P|\mu)$ for all realizable input costs μ. A joint distribution $H(P, \mu)$ summarizes both the collection of these conditionals together with an inherited distribution of unit input costs, $H(\mu)$.

Lemma 5 *Let a pdf $h(\mu)$ of a given distribution of the unit input cost exist and let also the assumptions of Proposition 1 be satisfied (a demand function $E[C|p]$*

is always a strictly decreasing function of price p). If $H(P, \mu)$ is a pricing strat-
egy maximizing the seller's profit, then the corresponding conditional strategies
$H(P|\mu)$ must be trivial distributions for all μ except for a set of measure zero,
i.e. any optimal pricing strategy can be represented by a function $\tilde{p}(\mu)$.

Proof: A pricing strategy affects $E[C|p]$ through an overall distribution of prices
only, regardless of what price is charged at what unit input cost. If the seller
modifies his pricing strategy in such a way that the resulting marginal distri-
bution of prices stays fixed, then the expectation $E[C|p]$ does not change. For
such a fixed price distribution:

$$
\begin{aligned}
E[\Pi] &= E\Big[E_c[C|p](p - \mu)\Big] \\
&= E_p\Big[E_c[C|p]p\Big] - E_\mu\Big[E_c[C|p]\mu\Big] \\
&= K - E_\mu\Big[E_c[C|p]\mu\Big],
\end{aligned}
\tag{58}
$$

where K is a constant depending on the marginal distribution of p only. The
maximal profit is achieved when the expectation of the total cost $E_\mu\Big[E[C|p]\mu\Big]$
is minimized, which is when low values of μ are aligned with high values of
$E[C|p]$ and vice versa. Since $E[C|p]$ is decreasing, then the highest prices must
be charged at the highest costs. We now show that the distributions $H(P|\mu)$
must be degenerate almost surely. If there is a positive measure of μ's (and thus
an uncountable number of these points) such that the distributions $H(P|\mu)$ are
supported by more than one point, then sets spanned by the supports of these
distributions must necessarily overlap. Low prices are not perfectly aligned
with lost costs. In such a case, the probability mass can be relocated to correct
this misalignment, while keeping both marginals fixed. The resulting pricing
strategy would lead to a higher profit. QED.

C   Numerics

Let the range of unit input costs be represented by N_μ cost points. A discretized
version of a pricing strategy is characterized by prices, $\{\tilde{p}_i\}_{i=1}^{N_\mu}$, at the cost points
$\{\mu_i\}_{i=1}^{N_\mu}$. Price values are allowed to vary continuously, $N_\mu = 10$. The expected
profit corresponding to a particular pricing strategy is evaluated via solving the

following discretized version of the consumer's problem $(5) - (8)$.

$$\{f_{ij}\}_{i=1,j=1}^{N_p,N_c} = \arg\max_{f'} \sum_{i=1,j=1}^{N_p,N_c} U(p_i,c_j)f'_{ij}, \tag{59}$$

subject to

$$\sum_{j=1}^{N_c} f'_{ij} = g_i, \qquad \forall i \tag{60}$$

$$f'_{ij} \geq 0, \qquad \forall i,j \tag{61}$$

$$\sum_{i=1,j=1}^{N_p,N_c} f'_{ij} \log \frac{f'_{ij}}{g_i(\sum_{k=1}^{N_p} f'_{kj})} \leq \kappa. \tag{62}$$

N_c is a number of consumption points and N_p a number of price points. This system is not difficult to solve numerically. (62) is a concave constraint defining a convex feasible set, while all other constraints as well as the objective are linear. Therefore, N_c can be quite large, $N_c = 300$, unlike N_μ. One can use any of the standard steepest-descent-based search algorithms, which are available even in R or Matlab. I use an optimization language AMPL together with a solver LOQO. It is easy to define the optimization problems in AMPL, while the computations are performed by the provided solver. LOQO applies interior-point methods, [13], which are efficient in solving large optimization problems of this type.

The seller's optimization task, however, is not concave; the problem might posses multiple local maxima. Therefore, some global optimization method has to be used. I chose a version of simulated annealing, which is a simple random search method. Any iteration of a pricing strategy leading to an increase in the seller's profit is accepted. On the other hand, unfavorable changes in profit are accepted with successively decreasing probability. The system "cools" down, and the upward shifts of the objective are gradually more and more preferred. This approach allows for escaping local maxima.

The algorithm consists of the following steps.

1. Initialize:

    (a) time $t := 0$,

    (b) initialize a pricing strategy, $\{\tilde{p}_i\}_{i=1}^{N_\mu}$,

(c) evaluate the profit, Π, corresponding to $\{\tilde{p}_i\}_{i=1}^{N_\mu}$ - solve the discretized version of the consumer's problem, (59)-(62), for $\{g_j\}_{j=1}^{N_p}$ generated by $\{\tilde{p}_i\}_{i=1}^{N_\mu}$; N_p is a number of different values of prices in $\{\tilde{p}_i\}_{i=1}^{N_\mu}$; $N_p \le N_\mu$.

2. Adjust time: $t := t + \Delta t$.

3. Adjust the system's temperature: $T(t) = at^{-b}$

   - typical values of the constant a, b in all computations: $a \in (10, 100), b \in (1, 3)$

   - simulated annealing often uses exponential cooling, power law however ensures better convergence results

4. Draw $\{\tilde{p}_i'\}_{i=1}^{N_\mu}$, a perturbation of $\{\tilde{p}_i\}_{i=1}^{N_\mu}$

   (a) draw k from $\{1..N_\mu\}$, specifying which price value is to be changed

   (b) draw x from a uniform distribution over $(0, 1)$, it determines a magnitude of a price change

   (c) perturb $\tilde{p}_k' := \tilde{p}_k + RT^{1/3}(2x - 1)$, where R is a constant related to a width of the range of optimal flexible prices

   (d) draw y from a uniform distribution over $(0, 1)$,
   if $y < 0.05$ and $k > 1$ then $\tilde{p}_k' = \tilde{p}_{k-1}$,
   if $y > 0.95$ and $k < N_\mu$ then $\tilde{p}_k' = \tilde{p}_{k+1}$,
   these two conditions ensure that situations with more prices collapsed on the same value are visited with a positive probability, prices are otherwise perturbed over a continuous range,

5. Evaluate the new profit Π' corresponding to the new strategy $\{\tilde{p}_i'\}_{i=1}^{N_\mu}$.

6. Let $test := \exp(\Pi' - \Pi)/(\Pi T)$.

7. Draw z from a uniform distribution over $(0, 1)$.

8. If $z \le test$, then adjust: $\{\tilde{p}_i\}_{i=1}^{N_\mu} := \{\tilde{p}_i'\}_{i=1}^{N_\mu}$ and $\Pi := \Pi'$.

9. If $t < t_{max}$, go to 2.

References

[1] Blinder, A.S: *Why are Prices Sticky? Preliminary Results from an Interview Study.* NBER Working Paper, 1991

[2] Eichenbaum, M., Jaimovich, N., Rebelo, S.: *Reference Prices and Nominal Rigidities.* 2008.

[3] Fabiani, S. et al.: *The Pricing Behavior of Firms on the Euro Area, New Survey Evidence.* European Central Bank, working paper, 2005.

[4] Mackowiak, B., Wiederholt, M.: *Optimal Sticky Prices under Rational Inattention.* unpublished, 2004.

[5] Mankiw N.G., Reis R.: *Sticky Information Versus Sticky Prices: A Proposal to Replace the New Keynesian Phillips Curve.* Quarterly Journal of Economics, 2002.

[6] Matějka, F.: *Rationally Inattentive Seller: Sales and Discrete Pricing.* working paper, 2008.

[7] Matějka, F., Sims, C.A.: *Discrete Actions in Information-Constrained Tracking Problems.* working paper, 2009.

[8] Nakamura, E., Steinsson, J.: *Price Setting in Forward-Looking Customer Markets.* 2005.

[9] Reis, R.: *Inattentive Producers.* Review of Economic Studies, 73, 2006.

[10] Sims, C.A.: *Stickiness.* Carnegie-rochester Conference Series on Public Policy, 49(1),1998.

[11] Sims, C.A.: *Implications of Rational Inattention.* Journal of Monetary Economics, 50, 2003.

[12] Sims, C.A.: *Rational Inattention: A Research Agenda.* 2005.

[13] Vanderbei, R.J.: *LOQO:an interior point code for quadratic programming.* Optimization Methods and Software, 11, 1999.

[14] Woodford, M.: *Information-Constrained State-Dependent Pricing.* unpublished.