Rigid Pricing and Rationally Inattentive Consumer[1]

Filip Matějka[2][3]

CERGE-EI[4]

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 the realized price, while the seller is perfectly attentive. The seller chooses to price rigidly, i.e. prices respond to changes in the unit input cost less than if the consumer were perfectly attentive. The rigidity is sometimes manifested by discreteness of prices - the seller charges a finite set of different prices even for a continuous range of unit input costs. The seller does so to provide the consumer with easily observable prices, which stimulates her to consume more and increases the seller's profit.

Keywords: rational inattention, imperfect information, nominal rigidity.

[1]This research was funded by GAČR P402/12/1993

[2]filip.matejka@cerge-ei.cz

[3]I am especially grateful to Per Krusell and Chris Sims for invaluable insights and guidance. I would also like to thank Jacques Cremer, Nobuhiro Kiyotaki, Alisdair McKay, Kristoffer Nimark, Ricardo Reis, Esteban Rossi-Hansberg, Jakub Steiner and Michael Woodford for helpful discussions and comments.

[4]A joint workplace of the Center for Economic Research and Graduate Education, Charles University, and the Economics Institute of the Academy of Sciences of the Czech Republic

1 Introduction

Developing theories of nominal rigidities and assessing those theories 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 explicit assumptions of an inability to alter prices (Calvo models). More recently, there has been considerable interest in the nominal rigidities implied by the price-setter's inability to process or acquire perfect information (Woodford, 2003; Reis, 2006). Furthermore, Maćkowiak and Wiederholt (2009), Woodford (2009) and Matějka (2010) study the pricing of rationally inattentive agents (Sims, 1998, 2003, 2006). All of these models have frictions microfounded on the price-setter' side. Prices do not change, either because they cannot, or it is costly to do so, or the price-setter does not know that the environment has changed.

This paper, on the other hand, shows that information friction in the form of rational inattention solely on the consumer's side can imply rigidity of prices, too. The motivation for focusing on frictions on the consumers' side is the explanation of nominal rigidities via the existence of implicit contracts between sellers and consumers, which is based on consumers' preferences for stable prices. Blinder (1991), and also Fabiani (2005) found through a series of interviews with price-setters that implicit contracts are likely to be one of the prominent reasons for prices being sticky. Gopinath and Rigobon (2008) also document the role of frictions on the consumers' side by showing that prices of imported goods are more sticky in the consumers' currency.

It seems plausible that consumers often possess less information about prices than sellers have about their input costs for instance - especially in markets where consumers need to compare the prices of several different sellers. When purchasing a product in a supermarket, many of us do not inspect all the prices in detail, our attention is limited. For instance, we can implicitly assume that prices end with .95 or .99, and if the number of cents is actually 85, we may not spot it and still keep our initial guess. Similarly, Lacetera et al. (2012) document consumers' inattention by the left digit bias in the used car market, where they find that consumers pay less attention to the right digits of millage on an odometer.

In the model presented here, the consumer processes information about the 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, which is stochastic.

The more flexible and thus more dispersed the pricing is, the more difficult it is for the consumer to observe the true realized price. If the consumer is more uncertain about the true value of the price, then she relocates a higher portion of her spending towards other products. This behavior is an analog to precautionary saving, where higher uncertainty about future shocks decreases current spending. Therefore, rigid pricing, which implies more precise prior knowledge about what price is realized, is particularly appealing. When the consumer processes little information, then prices are rigid; more specifically, the distribution of prices has a low entropy. In numerical solutions, 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, which generates a few price points only and thus a low entropy of prices. For consumers with a finite attention span, it is, for instance, easier to observe the price in the case that she knows a priori that it can take one of two values only. The seller benefits from a consumer having better knowledge about prices, which makes her consume more and which thus increases the seller's profit.

The presented model allows for full flexibility of rational information, which follows Sims (2006), while most of the literature on rational inattention assumes Gaussianity of uncertainty.[5] This generalization allows for the assessment of the informational complexity of pricing and the implication of lumpiness rather than simple rigidity of prices. With this general approach, the model in Matějka (2010) generates similar properties of pricing by assuming the information constraints on the price-setter's side instead; and Yang (2010) and Matějka and McKay (2012) apply it to an equilibrium setting of two players, too.[6]

The rest of the paper is organized as follows: The following section is devoted to the

[5] Luo and Young (2009); Maćkowiak and Wiederholt (2010); Van Nieuwerburgh and Veldkamp (2010); Mondria (2010); Paciello and Wiederholt (2011).

[6] Other papers using this formulation are Matějka and Sims (2010), Tutino (2013), Matějka and McKay (2014), and Stevens (2011).

formulation of the model. Section 3 presents a simplifying example which delivers rigid pricing when the seller is constrained to linear pricing strategies only, and Section 4 studies properties of solutions to the general formulation, which include the discreteness of pricing strategies.

2 The Model

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

2.1 Inattentive consumer

The consumer is rationally inattentive. She maximizes expectation of the indirect utility $U(p, c)$ from a consumption amount c when the price is p. She is endowed with prior knowledge about price given by the true distribution of prices $G \in \Delta(\mathbb{R})$, where $\Delta(\mathbb{R})$ is the set of all probability distributions on $\mathbb{R}$. The consumer's ability to process information about the realized price is limited.

The consumer chooses her information strategy $F_s \in \Delta_G(\mathbb{R}^2)$ and the action strategy $a : \Delta(\mathbb{R}) \to \mathbb{R}$. The information strategy F_s represents what the consumer pays attention to, and it is formalized as a joint distribution of prices and signals. $\Delta_G(\mathbb{R}^2)$ is the set of all probability distributions on $\mathbb{R}^2$ with the marginal distribution of price p being G, since F_s must be consistent with the prior G. This joint distribution fully describes what signals on the price are realized at a given price through the noisy process of information acquisition.

As the realized price is posted and the information about it is thus available, the choice of the joint distribution models the choice of attention to the price, or in other words the choice of the mental process of simplification and filtering of the available information. Allowing for signals of any form represents an unrestricted choice of such information filtering; the model abstracts from any additional assumptions of the physical availability of particular signals. The agent can choose to look at the price imperfectly in some unbiased way, e.g. with additive Gaussian noise; she can choose to use some form of partitioning of the price range, e.g. by reading the first digit of the price only, or she can use some sequential strategy of paying attention to more digits only in case the first one is low, etc.

The action strategy a describes how the agent acts upon the received signals; it maps a posterior belief $B \in \Delta(\mathbb{R})$ about the price to the optimal consumption level, such that $a(B) = \arg\max_{c\in\mathbb{R}} \mathbb{E}_B U(p, c)$. The action strategy thus defines a value function of each posterior belief B:

$$V(B) = \max_{c} \mathbb{E}_B U(p, c). \tag{1}$$

Since the consumer's ability to process information is limited, the concentration of the joint distribution F_s is constrained, too, i.e. there are bounds on co-movement of price p and signals s in the selected strategy. The information constraint (3) below, applied by using the findings of information theory, expresses a maximal achievable reduction of uncertainty about the price. This constraint was first used in Sims (2003); it can be motivated axiomatically as well as an achievable bound when the technology for information acquisition is from a specific fairly general class of technologies (Shannon, 1948; Cover and Thomas, 2006). What the agent can choose is the form of reduction of uncertainty. While the agent's ability to process information is limited, she is still completely free to process the pieces of information she cares about the most. In other words, the mutual information of F_s is constrained, but given the limits the consumer chooses how exactly it is shaped.

Definition 1. ***(The consumer's problem.)*** *Let $U(p, c)$ be the indirect utility function, $\kappa \geq 0$ be the consumer's information capacity, and G be the prior on p. The consumer's strategy $\{F_s, a\}$ is a solution to the following optimization problem.*

$$F_s = \arg \max_{F_s \in \Delta_G(\mathbb{R}^2)} \int_p \int_s V(F(\cdot|s)) F_s(ds|p) G(dp), \tag{2}$$

subject to

$$H[G] - E_s\left[H[F(p|s)]\right] \quad \leq \quad \kappa, \tag{3}$$

where a and V are given by the maximization in (1)*, and* $H[\cdot]$ *is the distribution's entropy.*

The integral in (2) is the consumer's expected utility, and (3) is the information constraint, which states that the expected entropy of posteriors must be equal or higher than the entropy of the prior less the information capacity κ. Entropy is a measure of uncertainty, it is a function of a distribution, and the constraint thus states what the maximal reduction of entropy is. For a distribution G with a pdf f the entropy is

$$H[G] = -\int_p f(p)\log f(p)dp,$$

and if the distribution is discrete, then the entropy is

$$H[G] = -\sum_k f_k \log f_k,$$

where f_k is the probability of state k.[7] For instance, the entropy of a Gaussian distribution of variance σ^2 is $1/2\log(2\pi e\sigma^2)$ and of a discrete distribution with N equally likely states it is $\log(N)$. According to (3), posteriors of lower entropy, and thus of less uncertainty, are less achievable.

The information-action strategy $\{F_s, a\}$ describes both the signals the consumer chooses to acquire as well as what c she selects upon receiving them. Receiving a signal s^* determines a specific posterior $F_s(p|s^*)$ about price, and then the consumer chooses consumption maximizing the expected utility, $c = a(F_s(p|s^*))$. However, in an optimal strategy, there is at most one signal generating a particular action c (see Matějka and McKay (2014) for the proof in an analogous setup). This is because otherwise the consumer would waste information on

[7] In general, it is $-\int_p f(p)\log f(p)\omega(dp)$, where ω is a base measure with respect to which the pdf exists. While the entropy is not invariant with respect to the choice of the base measure ω, the difference $H[G] - E_s\left[H[F(p|s)]\right]$ is.

distinguishing between two beliefs that generate the same outcome. Therefore, actions can be associated with signals and the induced joint distribution F of price and consumption describes signals, too.

Lemma 1. ***(The consumer's problem without signals.)*** *The joint distribution $F \in \Delta_G(\mathbb{R}^2)$ of price and consumption is induced by the solution to the problem in Definition 1 if, and only if, it solves the following optimization problem.*

$$F = \arg\max_F \int_p \int_c U(p,c)F(dpdc), \tag{4}$$

subject to

$$H[G] - E_c\left[H[F(p|c)]\right] \quad \leq \quad \kappa. \tag{5}$$

Proof: Matějka and McKay (2014).

We use this formulation of the consumer's problem throughout the rest of the paper. The conditional expectation of consumption given a price p, i.e. the demand curve, then takes the form:

$$E[c|p] = \int_c cF(dc|p). \tag{6}$$

The consumer would like to set a particular optimal consumption amount c for each price p. This would be represented by a degenerate joint distribution $F(p,c)$ located along the perfect-information demand curve only, which is not always possible since the average entropy of posteriors $F(p|c)$ must, according to the information constraint (5), be higher than $(H(G) - \kappa)$. The inattentive agent cannot observe the price, and is thus not able to set the optimal consumption amount perfectly.

2.2 Monopolistic seller

The monopolistic seller faces the rationally inattentive consumer. He incurs a random unit input cost μ drawn from $J \in \Delta(\mathbb{R})$ and supplies whatever amount is demanded by the consumer. The seller maximizes expectation of the profit Π.

$$\Pi \equiv c \cdot (p - \mu), \tag{7}$$

where c is the consumer's consumption, which equals the amount sold.

The seller chooses his pricing strategy $\tilde{p} : supp(J) \to \mathbb{R}$, which determines a unique price to be charged for a given drawn unit input cost μ, set price and commits to it.[8] $supp(J)$ denotes the support of J. The pricing strategy together with the distribution of μ determine the distribution of prices, which forms the consumer's prior G.

$$G(p) = \int_{\mu:\tilde{p}(\mu)<p} J(d\mu). \tag{8}$$

Each pricing strategy induces an optimal response by the consumer, and is thus also associated with a given level of expected profit which is determined by the resulting co-movement of prices and consumption in the consumer's strategy.

Definition 2. ***(The seller's problem.)*** *The pricing strategy* $\tilde{p} : supp(J) \to \mathbb{R}$ *maximizes the expected profit:*

$$\tilde{p} = \arg\max_{\tilde{p}'} \int E[c|\tilde{p}'(\mu)] \cdot (\tilde{p}'(\mu) - \mu) J(d\mu), \tag{9}$$

where the expected consumption, $E[c|\tilde{p}'(\mu)]$*, is given by* (6) *with* F *that solves the consumer's problem,* (4)-(5)*, for prior* G *that is given by* (8)*.*

If the consumer observed the price exactly, her consumption would be a function of the realized price only - and the seller could optimize $\tilde{p}$ on a cost by cost basis. However, when the consumer is constrained by a limited information capacity, consumption also depends on her prior knowledge about price, i.e. on the whole distribution of prices. The prior influences what forms of information the consumer chooses to acquire. The seller selects the whole pricing strategy simultaneously, considering the cross-effects of charging a specific

[8] For simplicity, the seller chooses from the set of pure strategies only. It is shown in Lemma 8 in Appendix B that even if mixed strategies were included in the set of admissible strategies, the optimal strategy would be a pure one.

price at one input cost on the consumer's responses to other prices charged at different input costs.

This model is a one-shot static Stackelberg problem. The seller commits to a pricing strategy, which he executes after the consumer picks her information-action strategy. Notice that if the seller could, it would always be optimal for him to ex post deviate from the strategy and charge higher prices. The model is supposed to represent a reduced form for a dynamic problem with reputation concerns. While the consumer does not observe prices, she acquires knowledge of the distribution from which the prices are drawn, and can punish the seller if the distribution is different than the one that the consumer expected. Moreover, taking the one-shot model literally, the consumer actually observes the price after each consumption decision, since she enjoys the resulting utility. For instance, in the following example, the consumer observes the amount of goods she consumes using what is left of the initial nominal endowment after the purchase from the seller.

3 Example

This section presents an illustrative example under the simplifying assumption that the seller selects only from among the linear pricing strategies. In this example, when the consumer has a low information capacity, then the seller chooses a pricing function that is less steep than the one under perfect information, i.e. prices are more rigid. In Section 4 the assumption of linearity of the pricing strategy is dropped, and while the rigidity of prices prevails, it can take a different form.

Let the indirect utility function be:

$$U_0(p,c) = 2c - c^2 + 2(1-pc) - (1-pc)^2. \tag{10}$$

The consumer consumes two goods, $U_0(p,c) = u(c) + u(1-pc)$, where $u(c) = 2c - c^2$ is the utility from consuming either of them. Her nominal endowment is 1, the price of good 1 is p, she chooses an amount c of good 1 to consume, while the remaining portion of her nominal endowment, $1 - pc$, is automatically used on consumption of good 2 which has the price 1. Given some posterior knowledge of price, the selected consumption amount is

$$c = \frac{1}{1 + E[p^2]} = \frac{1}{1 + E[p]^2 + \sigma_p^2}. \tag{11}$$

For a fixed expectation of price, the consumption decreases with the variance σ_p^2 of price in the posterior.

The seller chooses a linear strategy,

$$\tilde{p}(\mu) = p_0 + p_1\mu, \tag{12}$$

by specifying $p_0, p_1 \in \mathbb{R}$. Let μ be drawn from $N(0, \sigma_\mu^2)$ with σ_μ^2 being very small. The distribution of prices, and thus also the consumer's prior, is then $N(p_0, p_1^2\sigma_\mu^2)$. If the consumer observed the price, the optimal strategy as a solution to (9) would be:

$$\tilde{p}(\mu) = \mu + \sqrt{1 + \mu^2} = 1 + \mu + O(\mu^2) \tag{13}$$

The seller would select $p_0 = 1$ and $p_1 = 1$, and price would move one-to-one with the input cost.

In this example, the rationally inattentive consumer receives Gaussian signals on the price of the form $s = p_0 + p_1\mu + \epsilon$, where ϵ is the noise of variance that depends on the consumer's information capacity κ. The uncertainty $G(p)$ and $F(p|c)$ is Gaussian. It is known that in rational inattention the optimal signals under quadratic objectives and Gaussian uncertainty are Gaussian (Sims, 2003). Since σ_μ^2 is very small, then a linear-quadratic objective approximates the objective (10) well when the price is close to its ex-ante expectation and the assumed Gaussian signals approximate the optimal ones well, too. This is then a standard Bayesian updating under Gaussian uncertainty, which implies the following for the posterior expectation:

$$E[p] = p_0 + \xi p_1\mu + \sqrt{\xi(1-\xi)}p_1\eta, \tag{14}$$

where $\xi \in [0,1]$ scales responses to deviations in price, and $\sqrt{\xi(1-\xi)}p_1$ scales noise, let $\eta \sim N(0, \sigma_\mu^2)$. The parameter ξ is exogenously given, it depends on the relative precision of signals and the prior. It measures the refinement of knowledge, $(1-\xi)$ equals the ratio of the posterior and the prior variances of price. Therefore, ξ can be easily expressed as a monotone transformation of the information capacity κ, i.e. of the expected reduction of entropy between prior and posterior, without an explicit reference to the precision of signals.

$$\kappa = 1/2 \log \frac{1}{(1-\xi)}, \qquad \xi = 1 - e^{-2\kappa}, \tag{15}$$

since entropy of Gaussian variables of variance σ^2 equals $1/2\log(\pi e \sigma^2)$. $\xi = 0$ corresponds to $\kappa = 0$, and $\xi = 1$ to $\kappa = \infty$.

Now, plugging (14) into (11), we get:

$$c = \frac{1}{1 + (p_0 + \xi p_1 \mu + \sqrt{\xi(1-\xi)}p_1 \eta)^2 + (1-\xi)p_1^2 \sigma_\mu^2}. \tag{16}$$

The higher ξ the more information the consumer processes and the more $E[p]$ moves with the realized deviations in price, see (14). The distribution of p and $E[p]$ is jointly Gaussian, determining the joint distribution F of p and c in equation (4) via equation (16). The higher κ and ξ the stronger the co-movement of c and p.

To explore the effect of an increased flexibility of prices, i.e. higher p_1^2, on expected consumption, we use (16) and integrate it over shocks μ and noise η to get:

$$E[c] = \frac{1}{1+p_0^2} - p_1^2 \frac{\sigma_\mu^2}{(1+p_0^2)^2} \left(1 - \xi \frac{4p_0^2}{1+p_0^2}\right) + O((\sigma_\mu^2)^2). \tag{17}$$

First, with a fixed information capacity, higher p_1^2 increases the posterior uncertainty, which decreases consumption. This is the effect of $\left(-p_1^2 \frac{\sigma_\mu^2}{(1+p_0^2)^2}\right)$. On the other hand, increased variance of prices increases variance of $E[p]$ across realizations of prices and signals. If c is a convex function of $E[p]$, then this can increase expected consumption. This is the effect of $\left(-\xi \frac{4p_0^2}{1+p_0^2}\right)$. For low information capacity, i.e. low ξ, the first effect of decreased consumption dominates, $\left(1 - \xi \frac{4p_0^2}{1+p_0^2}\right)$ is positive. Higher p_1^2 decreases expected consumption when κ is low.

The analogous effect on the seller's profit follows, too. Maximization of expected profit for linear pricing strategies yields:

$$p_0 = \sqrt{\frac{3p_1^2\sigma_\mu^2}{2} + \frac{1}{2}, \sqrt{4 - 4p_1^2(\sigma_\mu^2)^2 + 9p_1^4(\sigma_\mu^2)^2}}, \qquad p_1 = \frac{(1+p_0^2)\xi}{1 + p_0^2(1-2\xi) + 2\xi}. \tag{18}$$

For $\xi = 0$, no information: $p_0 = 1$ and $p_1 = 0$. At zero information capacity, the price does not move with the input cost at all. Since p_0 changes continuously with ξ for small ξ, it can be expressed as $(1 + p_0'\xi + O(\xi^2))$ with a constant p_0'. Plugging this into (18) implies

$$p_1 = \xi + O(\xi^2),$$

p_1 increases with ξ at low ξ. Lower information capacity implies more rigid prices.

4 Solutions

Section 4.1 studies the properties of the solutions to the consumer's problem, (4) – (5). It discusses how the consumer's actions change when the price distribution changes, e.g. when it becomes more concentrated, while Section 4.2 discusses the resulting pricing strategies selected by the seller in (9). The difference from the example in Section 3 is that the seller is not constrained to linear strategies only, and the consumer then does not in general receive Gaussian signals.

4.1 Consumer's strategy

If the information constraint (5) is not binding, then the setup takes the form of a standard perfect information maximization problem. Conditional distributions of c given p degenerate at the perfect information demand function $c = c_{opt}(p)$.

$$c_{opt}(p) \quad = \quad \arg\max_{c'} U(p, c'), \tag{19}$$

In the rest of the section, we discuss properties of the consumer's strategy, when the information constraint (5) is binding. In this case, the analytical solution does not in general exist. The properties are the following. First, for a given price p, the consumption is dispersed about the optimal level $c_{opt}(p)$. Second, under some conditions, the consumer chooses to consume less if prices are more dispersed. Third, the consumer processes more information at prices where the potential losses from mis-perception are higher, which is typically at lower prices.

The first order condition on (4)-(5) implies $U(p, c) = \lambda \log(f(p|c)) + \xi(p)$, where $f(p|c)$ is the conditional pdf of price, λ is the Lagrange multiplier on (5), and $\xi(p)$ is the Lagrange multiplier on the consistency with the prior G, i.e. $\int_c F(p, dc) = G(p)$. This condition can then, using the consistency with prior, be rearranged to yield a description of consumption responses for a given price.

Lemma 2. ***(Probabilistic consumption response.)*** *For a given price p, the consumption response is probabilistic and follows the logit model.*

$$f(c|p) = \frac{f_c(c) e^{U(p,c)/\lambda}}{\int_{c'} e^{U(p,c')/\lambda} F_c(dc')}, \tag{20}$$

where $f(c|p)$ is the conditional pdf of consumption and f_c, resp. F_c, is the unconditional pdf, resp. cdf, of consumption.

The consumer selects a suboptimal level of consumption, i.e. makes mistakes, due to imperfect information about the price. The mistakes, however, have a particular form expressed by (20). For a given price, the likelihood that the consumption c is selected increases with $U(p, c)$. The higher the information capacity κ, the lower is the Lagrange multiplier λ, and the distribution of consumption is more concentrated in the regions of high $U(p, c)$, which is around $c_{opt}(p)$. Notice that the logit-based formula states that the likelihood that the consumption c is selected also increases with the ex ante unconditional probability that c is selected, which reflects the Bayesian updating.[9] The higher κ, the weaker the effect of prior beliefs is on the choice probabilities and the more important is $U(p, c)$. In the limit, as $\kappa \to \infty$ and $\lambda \to 0$, the optimal consumption $c_{opt}(p)$ is selected with probability one. The lower κ is, the more the induced demand curve $E[c|p]$ deviates from the optimal demand curve.[10]

Figure 1 shows numerical solutions for the utility function U_0, price uniformly distributed in $(2, 3)$, and three levels of information capacity κ.[11] Each joint distribution is presented in two different ways, one more suitable for presenting the support of the distribution and one for its shape. The dashed curve in the upper exhibits is the optimal demand function, $c_{opt}(p)$. As the information capacity increases, the distribution is more tightly concentrated about $c_{opt}(p)$. This is what the first order condition (20) suggests.

Notice that the consumer uses an information strategy that has features of a partitioning. She does not receive signals with additive noise, for instance, but rather she distinguishes, albeit imprecisely, whether the price is in particular ranges. It is clearly visible that the consumer decides to process a finite number of realizations of signals only, the higher the capacity the higher this number is. The discrete signals then lead to a finite number of different consumption levels only; two levels for $\kappa = 0.5$, three for $\kappa = 1$ and seven for $\kappa = 2$. The unconditional distribution F_c is discrete.[12] For instance, for $\kappa = 0.5$ prices close to 2 most often generate consumption equal to 0.22, while prices close to 3 generate $c = 0.11$ with high probability. One signal reflects that the price is most likely below 2.4, while the other reflects that price is most likely above 2.4.

Let us now explore the effect of dispersed beliefs on the level of consumption.

[9]See Matějka and McKay (2014) for a detailed study of the connection between rational inattention and the logit model.

[10]Lemma 7 in Appendix A states conditions under which both $c_{opt}(p)$ and $E[c|p]$ are strictly decreasing functions of p.

[11]The original problem (4)-(5) was discretized into the corresponding finite dimensional convex problem. Details of the numerical method are described in Appendix C

[12]The discreteness of actions of rationally inattentive agents is explored in Matějka and Sims (2010) and Matějka (2010)

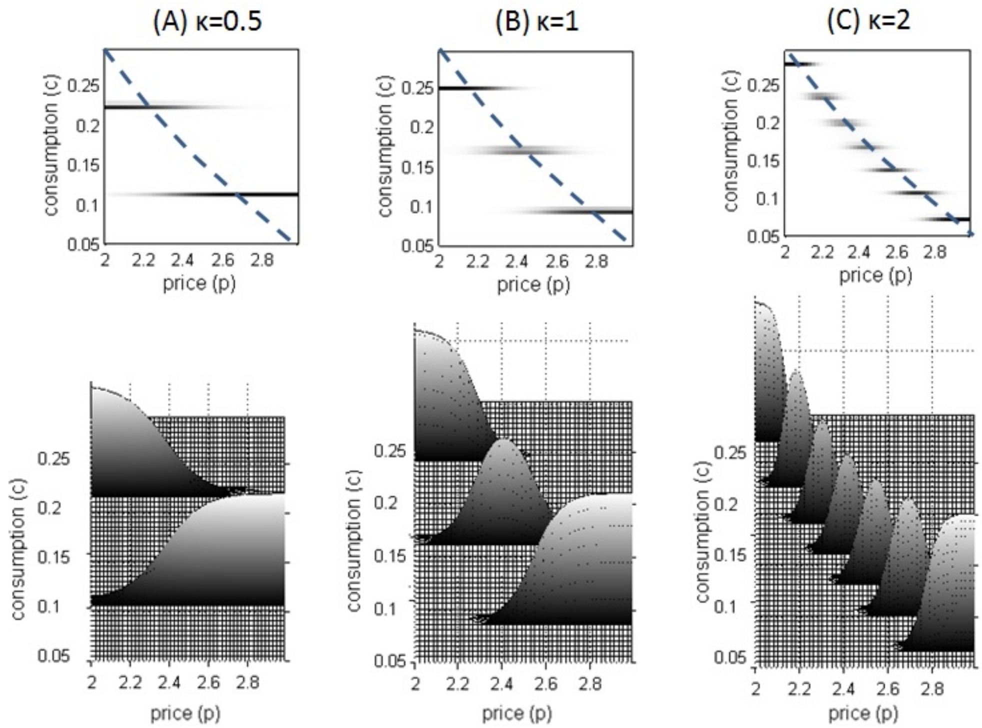

Figure 1: Joint distributions of prices and consumption, U_0

Lemma 3. ***(Effect of imperfect information on consumption.)*** *Let c^* be the optimal level of consumption for a belief $B \in \Delta(\mathbb{R})$ about the price, and c^{**} be the optimal consumption for belief $B' \in \Delta(\mathbb{R})$, which is a mean preserving spread of B. If $U(p,c)$ is concave in c, and $\frac{\partial U(p,c)}{\partial c}$ exists and is strictly concave in p, then*

$$c^* > c^{**}.$$

Proof: Appendix A.

More dispersed belief about price of the same expectation leads to lower consumption. The lemma implies that volatile prices drive consumption down. This is because with a fixed information capacity, more dispersed prices (and thus more dispersed prior belief) imply less precise posterior beliefs about the price. Consumers react to unstable prices by consuming less. In the two-good interpretation of the objective U_0, the consumer fears that she loses utility from consuming too little of good 2, in case the seller's price of good 1 turns out to be unexpectedly high. If the price p is volatile, she rather consumes more of good 2 with the known price.

This effect is an analog to precautionary saving in a standard two period savings problem with a stochastic stream of endowments. There a consumer maximizes $\hat{u}(c_1) + E[\hat{u}(y - c_1)]$, where $\hat{u}(\cdot)$ is the one-period utility from consumption, c_1 is the selected consumption in the first period, and y is the uncertain permanent income. In the savings problem, the first order condition is:

$$\hat{u}'(c_1) = E\left[\hat{u}'(y - c_1)\right], \tag{21}$$

The equation (21) implies that if $\hat{u}$ is concave and $\hat{u}'$ convex, then the selected c_1 decreases when we replace the random variable y with $y' = y + \epsilon$ such that $E[\epsilon|y] = 0$, i.e. when y undergoes a mean preserving spread and thus becomes more uncertain. If uncertainty about future endowments increases, precautionary saving increases, too.

In our model, let the utility function $U(p, c)$ have a two-good interpretation, just like U_0, but for a general u:

$$U(p, c) = u(c) + u(1 - pc),$$

where $u(\cdot)$ is utility from consuming either good. The role of future consumption amounts from the consumption-savings problem is played in our model by the consumption of good 2. The inattentive consumer maximizes $u(c) + E[u(1 - pc)]$. For a given posterior belief about p, the first order condition, analogous to (21), is:

$$u'(c) = E\left[pu'(1 - pc)\right]. \tag{22}$$

Uncertainty about the amount of consumption of good 2 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. Similarly to the consumption-savings problem, here if u is concave and $pu'(1 - pc)$ is convex in p, then the consumption c of the seller's good decreases when uncertainty increases. It is easy to check that this is exactly if $U(p, c) = u(c) + u(1 - pc)$ satisfies the assumption in Lemma 3.

The assumptions on $U(p, c)$ that are needed to generate this effect do not seem to be very restrictive. For instance, any utility function with the two-good interpretation (4.1) and the standard properties $u'' < 0$ and $u''' > 0$ satisfies the assumptions. Consider also two other utility functions, a CES aggregator, or its limit version as the share parameter a approaches zero.

$$U_1(p, c) = \Big(a \cdot c^r + (1 - a) \cdot (1 - pc)^r\Big)^{1/r}, \qquad U_2(p, c) = c^r - pc, \tag{23}$$

where $r = 1 - 1/\theta$ is the elasticity of substitution, $\theta \in (1, \infty)$. Again, the consumer derives utility from consuming the seller's product and also from consuming some other good of price 1. After purchasing from the seller, she automatically spends all that is left of the initial unit nominal endowment on purchasing the second good. U_1 satisfies the assumptions, but U_2 does not, since in the limit $a \to 0$, the effect of c on disutility from consuming good 2 is small and thus linear. While for U_0 and also for U_1 the demand decreases with uncertainty about the price, see for instance (11), for U_2 the consumption amount is proportional to $E[p]^{-\theta}$, and it is independent of any higher moments.

Finally, the consumer chooses what information to process based on the relative importance of various pieces of information given by the form of her utility function. The first order condition (20) implies that the conditional distribution of consumption $f(c|p)$ is more concentrated for prices p, for which $U(p, c)$ varies more with c, i.e. when the stakes from suboptimal actions are higher. The consumer desires to reduce the part of uncertainty that would otherwise induce the largest utility losses.

Figure 1 shows that the consumer chooses to acquire slightly tighter signals when the price is low. Using the Taylor expansion, the loss in utility due to misjudging a price p by a small amount ϵ is approximately equal to $L(p)\epsilon^2/2$, where the loss factor $L(p)$ is:

$$L(p) = \left| \left(\frac{dc_{opt}(p')}{dp'} \right)^2 \frac{d^2U(p, c)}{dc^2} \right|, \qquad \text{at } c = c_{opt}(p), p' = p. \tag{24}$$

The higher the loss factor, the more attention the consumer pays to that region of p. The loss factors are decreasing for U_0 as well as both for U_1 and U_2,[13] and it takes an especially simple form for U_2:

$$L(p) \propto p^{-\theta-1}. \tag{25}$$

The potential losses are higher at lower prices. This finding that the consumer pays more attention to low prices is likely to hold more generally, since at low prices the consumer purchases a high amount and thus any misjudging of the price is particularly costly.

4.2 Seller's pricing strategies

By modifying the pricing strategy and thus the distribution of prices, too, the seller incentivizes the consumer to process different pieces of information, attain different posterior beliefs and thus respond differently, potentially even to the same realized and imperfectly observed true price. We discuss below that the optimal strategies have the following properties. First, the prices are rigid, i.e. they do not move perfectly with the input cost. Second, lower prices are more flexible than the higher ones.

When $\kappa = \infty$, then the consumer knows the realized price exactly. Her demand for the good depends on the realized price only, not on the whole distribution of prices. The seller sets an optimal price for each cost separately, he does not need to consider the implications of the prior for the form of the posterior. For instance, for U_0 the optimal pricing strategies are given by (13).

When $\kappa = 0$, on the other hand, the consumer cannot 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. Optimal prices for different input costs are not set independently of each other.

The following lemma is an implication of Lemma 3. If the consumer does not process any information, $\kappa = 0$, then her posterior belief equals the distribution of prices, and the

[13]The loss factor for U_2 is in fact decreasing for $p > 1$ only, which is not restrictive, because the seller's price is in equilibrium greater than 1 anytime $\mu \geq 0$.

consumption is given by this distribution, too, regardless of the realized price. The seller's profit then equals the expected markup, which is the expected price less the expected unit input cost, times the consumption. However, if the distribution of prices were non-trivial, then Lemma 3 implies that consumption increases if the price were instead kept fixed at its expectation, and the profit would thus increase, too.

Lemma 4. ***(Constant pricing for*** $\kappa = 0$***.)*** *If the assumptions of Lemma 3 are satisfied, then for* $\kappa = 0$ *the optimal* $\tilde{p}$ *must be constant - the seller charges the same price for all unit input costs.*

Proof: Appendix B.

Next, we address positive information capacities. If $\kappa > 0$ then the consumer refines the prior knowledge and the seller can find it beneficial to charge different prices at different input costs. However, the implication of Lemma 3 still applies. The more dispersed the prices are, the more dispersed beliefs are, and the less the consumer consumes, which has a negative impact on profit.

Proposition 1. ***(Rigid pricing for small*** κ***.)*** *If the assumptions of Lemma 3 are satisfied, then for any non-constant pricing strategy* $\tilde{p}$*, there exist* κ_0 *such that if* $\kappa < \kappa_0$*, then* $\tilde{p}$ *is not the optimal pricing strategy. More specifically, there exist* $\delta > 0$ *and* p_0 *such that for all* $\kappa < \kappa_0$*, all pricing strategies* $\tilde{p}'$ *that satisfy* $|\tilde{p}' - p_0| < \delta$*, i.e. they are sufficiently close to a constant strategy, generate a higher profit than* $\tilde{p}$ *generates, and also there exist pricing strategies with a finite number of prices of entropy equal to* κ *that generate a higher profit than* $\tilde{p}$ *does.*

Proof: Appendix B.

The proposition relates arbitrarily rigid prices, i.e. within a small interval or attaining a small number of values, if the consumer is sufficiently inattentive, yet not necessarily completely inattentive. First, when prices are continuous, then they must be close to a constant price to be of low entropy. For any distribution with a pdf, entropy is proportional to the log of the variance. As the distribution's variance or width is scaled, all states become

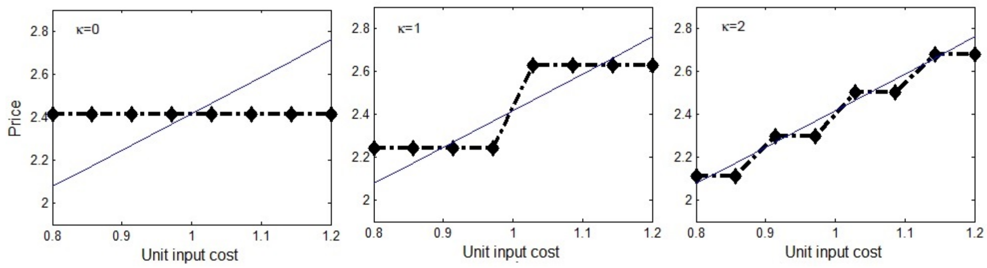

Figure 2: Pricing strategies, U_0.

less likely, and entropy increases, which implies more posterior uncertainty, and thus lower consumption, too. Second, however, if the distribution of prices is discrete, then simply moving the price-points further apart does not change their entropy, while it increases the variance. For instance, if there are two price-points, each realized with the probability 0.5, then their entropy is 1, i.e. the consumer needs to process 1 bit of information, regardless of the distance between the two prices. In the continuous case, spreading the distribution makes it thinner and single prices become less likely, which does not happen in the discrete case, when their probabilities stay fixed. This is why in the discrete case, spread-out prices can be optimal even when κ is low.

There are 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 tends to choose more condensed pricing, realizing that consumption falls when the signal on price is more dispersed, which is stated in Lemma 3. 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 set different prices 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 for low κ.

Let us now explore solutions numerically. Figure 2 shows numerical solutions for U_0 and different levels of information capacity. The solid line in each graph represents the perfect-information pricing strategy (13), $\kappa = \infty$. The higher the information capacity, the closer the seller's strategy and the perfect-information strategy are.

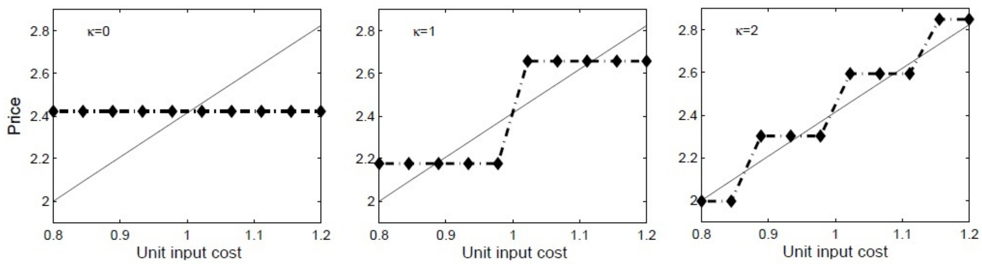

Figure 3: Pricing strategies, U_1, $\theta = 2$, $a = 0.5$

The graph on the left, when the consumer is completely inattentive, is a manifestation of Lemma 4 - the optimal pricing strategy is constant and the price is completely rigid. If the information capacity increases, pricing becomes more flexible, entropy of prices increases, which is related to the statement in Proposition 1. 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. Figure 3 shows very similar numerical results for the utility function $U_1, \theta = 2$, and $a = 0.5$. In all cases in the two figures, the seller chooses a strategy of maximal entropy such that the consumer can still observe the price exactly. 1 bit of information capacity allows the consumer to distinguish between two different values, while 2 bits distinguish between four of them.

Optimal pricing strategy is not, however, in general such that the prices can be observed. Figure 4 shows optimal pricing in a case with a lower precautionary motive; it is for U_1, $\theta = 2$ but with a smaller share parameter of the good 1, $a = 0.25$, resp. $a = 0.1$. The utility function becomes more similar to U_2, the precautionary motive reflected by $|\partial^3 U(p, c)/\partial c \partial p^2|$ decreases and prices become more flexible. There are three, resp. five, levels of prices, which cannot for this distribution be distinguished with $\kappa = 1$. Notice also that prices first become more flexible at lower input costs. This is where, according to the previous section, the consumer chooses to pay more attention, and thus where the consumer's uncertainty about the price is better resolved.

While Proposition 1 suggests that at sufficiently low κ the equilibrium prices are either of low variance or discrete, the numerical solutions reveal a tendency towards discreteness. I was not able to prove the existence of discreteness analytically, given the generality of

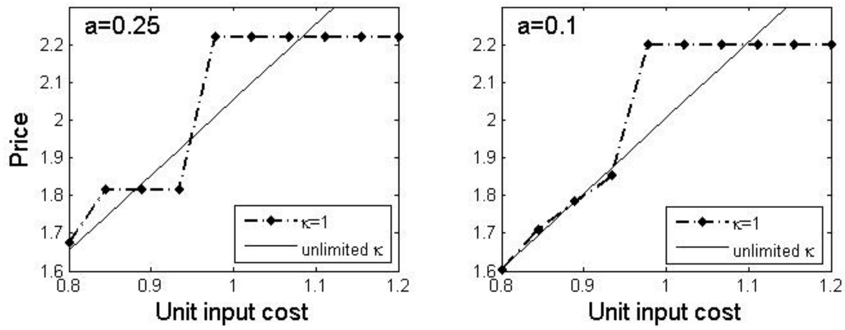

Figure 4: Pricing strategies, U_1, $\theta = 2$

information structures that the consumer can acquire, yet we can discuss what the forces are that could generate discreteness.

First, let us recall the two forces shaping the optimal pricing strategies that we discussed above. If, for some reason, the first force dominates, which is that prices are concentrated due to precautionary motives to consume less under imperfect information, then the seller wants the consumer to observe the price. It is possible that this force dominates at low κ, as then the second force, the motive to charge different prices at different marginal costs, diminishes as prices are not distinguished. However, observing the price for $\kappa < \infty$ is possible only if the distribution of prices is discrete. Anytime the distribution is continuous then the consumer would need an infinite information capacity to observe the price, since entropy of the delta function is minus infinity. Discrete prices allow for perfect posterior knowledge even with finite κ, e.g. anytime $\kappa \geq \log(N)$, where N is the number of price-points. This might be the case in Figures 2 and 3, where the optimal pricing is such that the consumer observes the price.

Second, Figure 1 suggests that consumers find it optimal to use information strategies similar to partition-based ones. They do not receive signals with additive noise, but rather they find out whether the price is in a certain range, e.g. for $\kappa = 0.5$ the consumer distinguishes between two beliefs only. While the consumer's strategy is not exactly partition-based, the close similarity provides force towards discreteness.

Let us call an information strategy "partition-based" if there is a countable set of nonoverlapping intervals I_n such that $\bigcup_{n=-\infty}^{\infty} I_n = \mathbb{R}$, and a countable set of corresponding signals $\{s_n\}_{n=-\infty}^{\infty} \subset \mathbb{R}$ such that for $p \in I_n$ the consumer receives the signal s_n with probability one, i.e. she observes what interval I_n the price belongs to. If the consumer used a partition-based strategy, then prices would always be discrete.

Corollary 1. ***(Discreteness for partition strategies.)*** *If the assumptions of Lemma 3 are satisfied and if the consumer uses a fixed partition-based information strategy, then the set of prices is discrete.*

Proof: This is an immediate implication of Lemma 4 applied to each interval I_n separately. The optimal pricing strategy must be constant on each I_n since the consumer does not refine his knowledge beyond receiving s_n, i.e. which interval the price belongs to. $\square$

Notice that the partition-based strategies do not necessarily imply that prices would be placed at the right edges of each interval, because the consumer knows the unconditional distribution, but only that there is at most one price-point within each interval.

To conclude this section, Figures 5 and 6 present simulated trajectories of price if the static model were repeated over several periods. This setup corresponds, for instance, to the case where the seller faces a continuum of consumers visiting the seller's store in different periods. Each of the consumers visits the store only once, which makes the problem static from the consumer's point of view and thus from the seller's too, since he does not face any additional frictions of his own. What generates the consumer's prior is the long-term distribution of μ and the pricing strategy.

A simulated price series corresponding to the pricing strategy depicted in the upper-right corner in Figure 3 is shown in Figure 5. The figure compares optimal pricing when the 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. We simply generated a serially correlated μ from a stochastic process with a long-term distribution that is uniform in $(0.8, 1.2)$ and then applied the pricing strategy repeatedly.

Figure 6 shows a simulated time series of prices for $a = 0.25$ - the corresponding optimal

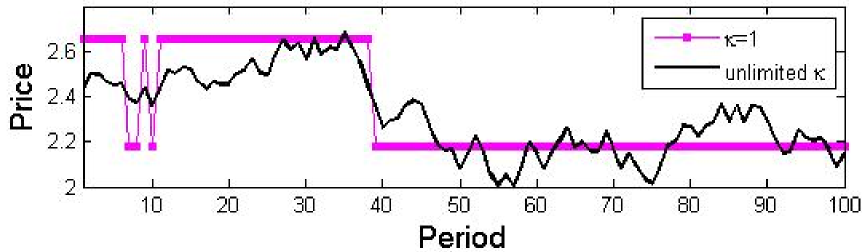

Figure 5: Simulated price series, U_1, $\theta = 2$, $a = 0.5$.

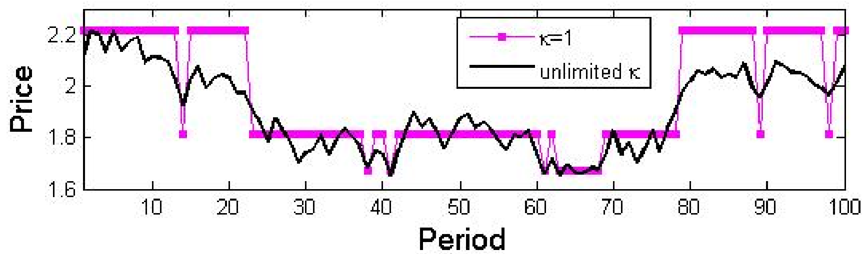

Figure 6: Simulated price series, U_1, $\theta = 2$, $a = 0.25$.

pricing strategy is on the left in Figure 4. The lower share parameter a implies more flexible pricing; prices take 3 different values. The lower prices, i.e. sales, are more flexible and transient.

5 Conclusion

This paper proposes a mechanism leading to rigid pricing as the price-setter's optimal response to a consumer's limited attention to prices. The rigidity can, under the general formulation of rational inattention, take the form of prices moving continuously but less than one-for-one with input costs, or the pricing strategy can be discrete. 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.

While the information constraint on the consumer's side has, I believe, intuitive appeal and also some empirical support (Blinder, 1991; Fabiani, 2005; Gopinath and Rigobon, 2008), the interaction between a price-setter and a rationally inattentive consumer makes it difficult to solve the model in a more complex setting.

The next step would be to find a way to embed this mechanism in a larger dynamic model, which could ultimately be used to study optimal policies in a model with rationally inattentive consumers. Paciello and Wiederholt (2011) find that when price-setters are rationally inattentive, then stabilizing-price level is particularly appealing, more than in models with Calvo-pricing. At first sight, it seems that the model presented here could have similar implications. Finally, empirical research more directly targeted at exploring the relevance of the information constraints assumed here and at the connection between inattention and price rigidity would be very useful, too.

References

Aarts, E. H. and Korst, J. (1989). Simulated annealing and boltzmann machines. a stochastic approach to combinatorial optimization. *Chichester, New York: Wiley, 1989*, 1.

Blinder, A. S. (1991). Why are prices sticky? preliminary results from an interview study. *American Economic Review*, 81(2):89–96.

Cover, T. M. and Thomas, J. A. (2006). *Elements of Information Theory*. Wiley, Hoboken, NJ.

Fabiani, S. (2005). The pricing behaviour of firms in the euro area: New survey evidence. Technical report.

Gopinath, G. and Rigobon, R. (2008). Sticky borders. *The Quarterly Journal of Economics*, 123(2):531–575.

Lacetera, N., Pope, D. G., and Sydnor, J. R. (2012). Heuristic thinking and limited attention in the car market. *The American Economic Review*, 102(5):2206–2236.

Luo, Y. and Young, E. R. (2009). Rational inattention and aggregate fluctuations. *The B.E. Journal of Macroeconomics*, 9(1).

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

Maćkowiak, B. A. and Wiederholt, M. (2010). Business cycle dynamics under rational inattention. CEPR Discussion Papers 7691, C.E.P.R. Discussion Papers.

Matějka, F. (2010). Rationally inattentive seller: Sales and discrete pricing. CERGE-EI, Working Paper 408.

Matějka, F. and McKay, A. (2012). Simple market equilibria with rationally inattentive consumers. *The American Economic Review*, 102(3):24–29.

Matějka, F. and McKay, A. (2014). Rational inattention to discrete choices: A new foundation for the multinomial logit model. Unpublished.

Matějka, F. and Sims, C. A. (2010). Discrete actions in information-constrained tracking problems. Technical report, Princeton University.

Mondria, J. (2010). Portfolio choice, attention allocation, and price comovement. *Journal of Economic Theory*, 145(5):1837 – 1864.

Paciello, L. and Wiederholt, M. (2011). Exogenous information, endogenous information and optimal monetary policy. Working paper, Northwestern University.

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

Shannon, C. (1948). A mathematical theory of communication. *Bell System Technical Journal*, 27.

Sims, C. A. (1998). Stickiness. *Carnegie-Rochester Conference Series on Public Policy*, 49(1):317–356.

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

Sims, C. A. (2006). Rational inattention: Beyond the linear-quadratic case. *American Economic Review*, 96(2):158–163.

Stevens, L. (2011). Price adjustment in a model with multiple price policies. *Working paper, Columbia University.*

Tutino, A. (2013). Rationally inattentive consumption choices. *Review of Economic Dynamics*, 16(3):421–439.

Van Nieuwerburgh, S. and Veldkamp, L. (2010). Information acquisition and under-diversification. *Review of Economic Studies*, 77(2):779–805.

Vanderbei, R. J. (1999). Loqo: An interior point code for quadratic programming. *Optimization methods and software*, 11(1-4):451–484.

Woodford, M. (2003). Imperfect common knowledge and the effects of monetary policy. *Knowledge, Information, and Expectations in Modern Macroeconomics: In Honor of Edmund Strother Phelps*, page 25.

Woodford, M. (2009). Information-constrained state-dependent pricing. *Journal of Monetary Economics*, 56:S100–S124.

Yang, M. (2010). Coordination with rational inattention. Unpublished.

A Proofs concerning consumer's strategies

Proof of Lemma 3. The optimal response c^* to belief B satisfies

$$E_B\left[\frac{\partial U(p,c)}{\partial c}\right]_{c=c^*} = 0. \tag{26}$$

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

$$E_{B'}\left[\frac{\partial U(p,c)}{\partial c}\right]_{c=c^*} < E_B\left[\frac{\partial U(p,c)}{\partial c}\right]_{c=c^*} = 0, \tag{27}$$

which implies:

$$c^* > c^{**}, \tag{28}$$

where c^{**} is such that $E_{B'}\left[\frac{\partial U(p,c)}{\partial c}\right]_{c=c^{**}} = 0.$ □

Definition 3. *Let $h(c)$ be a marginal pdf of c with respect to a probability measure ω_c. A pair of functions $\{F_1(c), F_2(c)\}$ has the* m-property*, 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 5. *Let $f(c|p)$ be a conditional pdf with respect to ω_c of a solution to* (4)-(5) *and let $p_1 < p_2$ be in the support of the prior G. 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. The first order condition, expressed in (20) can also be written as:

$$f(c|p) = h(c)\exp(U(p,c)/\lambda)w(p), \tag{29}$$

where $h(c)$ is the marginal pdf of c, $\lambda > 0$ is a Lagrange multiplier on the information constraint, (5), and $w(p)$ is a normalization parameter. If $p \in supp(G)$, 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,

$$\exp(U(p_1, c)/\lambda)w(p_1) - \exp(U(p_2, c)/\lambda)w(p_2) \tag{30}$$

also has at most one zero and is of the opposite sign on either side of some value m for c. But from (29) 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. □

Lemma 6. *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.

$$\begin{aligned}
E[C|p_1] - E[C|p_2] &= \int_{c<m} (f(c|p_1) - f(c|p_2))\, c\omega_c(dc) + \int_{c>m} (f(c|p_1) - f(c|p_2))\, c\omega_c(dc) \geq \\
&\geq \int_{c<m} (f(c|p_1) - f(c|p_2))\, m\omega_c(dc) + \int_{c>m} (f(c|p_2) - f(c|p_1))\, c\omega_c(dc) = \\
&= -\int_{c>m} (f(c|p_1) - f(c|p_2))\, m\omega_c(dc) + \int_{c>m} (f(c|p_1) - f(c|p_2))\, c\omega_c(dc) \geq \\
&= \int_{c>m} (f(c|p_1) - f(c|p_2))\, (c - m)\omega_c(dc) \geq 0.
\end{aligned} \tag{31}$$

Both inequalities hold assuming the m-property. In the last equality, we used the fact 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. □

Lemma 7. *Let $\frac{\partial^2 U}{\partial c \partial p} < 0$ and $\kappa > 0$, then $c_{opt}(p)$ as well as $E[c|p]$ are strictly decreasing functions of p in the support of the prior on p.*

Proof. This is an immediate consequence of Lemmata 5 and 6. □

B Proofs concerning seller's strategies

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 described by a joint distributions $S(P, \mu) \in \Delta(\mathbb{R}^2)$. The following lemma justifies why the seller in the presented model optimizes over pure strategies only.

Lemma 8. *Let the assumptions of Lemma 7 be satisfied (a demand function $E[c|p]$ is thus a strictly decreasing function of p), then the conditional strategies $S(p|\mu)$ must be trivial distributions for all μ except for a set of μ's of zero measure, i.e. any optimal pricing strategy can be represented by a function $\tilde{p}(\mu)$ almost surely.*

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 distribution of prices stays fixed, then the expectation $E[c|p]$ does not change. For such a fixed price distribution:

$$\begin{aligned} E[\Pi] &= E\Big[E_c[c|p](p-\mu)\Big] \\ &= E_p\Big[E_c[c|p]p\Big] - E_\mu\Big[E_c[c|p]\mu\Big] \\ &= K - E_\mu\Big[E_c[c|p]\mu\Big], \end{aligned} \tag{32}$$

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 $S(p|\mu)$ 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 $S(p|\mu)$ 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. □

Proof of Lemma 4. Let us assume that a non-constant $\tilde{p}_1$ is an optimal strategy. Such a strategy delivers a non-degenerate distribution of prices. Since the consumer has zero information capacity, her posterior knowledge equals her prior. 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

$$c < c'. \tag{33}$$

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

$$\begin{aligned} E_{\tilde{p}_1}[\Pi] &= c \cdot E_{\tilde{p}_1}[p - \mu] < \\ &< c' \cdot E_{\tilde{p}_1}[p - \mu] = c' \cdot (\bar{p} - \bar{\mu}) = E_{\tilde{p}_2}[\Pi], \end{aligned} \tag{34}$$

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

Proof of Proposition 1. First, Lemma 3 implies:

$$c_G < c_{opt}(E_G[p]),$$

where c_G is consumption given a posterior belief that equals the prior G, and $c_{opt}(E_G[p])$ is the consumption under perfect information if the price equals expectation of G.

Second, for all $\epsilon_1 > 0$ there exists $\kappa_1 > 0$ such that for all $\kappa < \kappa_1$ and all realized consumption levels c:

$$|c - c_G| < \epsilon_1.$$

If information capacity κ is sufficiently small, then the posterior beliefs are sufficiently close to the prior G, and thus also consumption levels are close to c_G.

These two points imply that for a given non-trivial prior of a given entropy higher than

$H[\delta]$, there exists a small enough information capacity κ_0 such that for all $\kappa < \kappa_0$

$$c < c_{opt}(E_G[p]).$$

Similarly as in the proof of Lemma 4, this implies that for the given pricing strategy the profit Π is lower than if the strategy were a constant price equal to $E_G[p]$:

$$\Pi = E\left[E[c|p](p(\mu) - \mu)\right] < c_{opt}(E_G[p])E\left[(p(\mu) - \mu)\right] = c_{opt}(E_G[p])\left(E_G[p] - E_J[\mu]\right).$$

For each pricing strategy and such non-trivial G, there exists $\kappa_0 > 0$ such that the strategy cannot be optimal for any $\kappa < \kappa_0$.

Now, since the constant pricing strategy generates a higher profit than $\tilde{p}$, and since the problem is continuous, all pricing strategies $\tilde{p}'$ in a sufficiently small neighborhood of the constant strategy, i.e. such that $|\tilde{p}' - E_G[p]| < \epsilon_2$, generate a higher profit, too.

Moreover, there exist discrete strategies with the entropy of G equal to κ that also generate a higher profit. Such strategies imply that the price is observed by the consumer, since the entropy of posteriors equals $(H[G]-\kappa) = 0$. The support of the distribution of input costs J can then be partitioned into intervals such that the entropy of the intervals equals κ, i.e. entropy of the discrete distribution given by probability masses of the intervals equals κ. Then, let the pricing strategy be such that it is constant within each of the intervals, equal to the optimal perfect information price when the unit input cost equals the expected unit input cost conditional on the interval. Such prices will in general be non-constant, and since they are set optimally (the prices are observed) and the constant price across the intervals is feasible, too, they generate a higher profit than the constant price, and thus also than the original strategy $\tilde{p}$. □

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_μ is finite and small. The expected profit corresponding to a particular pricing strategy is evaluated via solving the following discretized version of the consumer's problem (4)-(5).

$$\{f_{ij}\}_{i=1,j=1}^{N_p,N_c} = \arg\max_{f'} \sum_{i=1,j=1}^{N_p,N_c} U(p_i, c_j) f'_{ij}, \tag{35}$$

subject to

$$\begin{aligned} \sum_{j=1}^{N_c} f'_{ij} &= g_i, \qquad \forall i & (36)\\ f'_{ij} &\geq 0, \qquad \forall i, j & (37)\\ \sum_{i=1,j=1}^{N_p,N_c} f'_{ij} \log \frac{f'_{ij}}{g_i(\sum_{k=1}^{N_p} f'_{kj})} &\leq \kappa. & (38) \end{aligned}$$

N_c is a number of consumption points and N_p a number of price points. This system is not difficult to solve numerically. (38) 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, Vanderbei (1999), 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(Aarts and Korst, 1989), 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.