Actively Learning by Pricing: A Model of an Experimenting Seller*

Tim Willems†

September 13, 2011

Abstract

This paper presents a model of a monopolist who is actively learning the time-varying slope of his demand curve through his pricing strategy. This provides the seller with an incentive to experiment with his price. In particular, the optimal pricing strategy prescribes the alternation of periods in which a high price is quoted with occasional mark-downs of a rather fixed size. Consequently, the model can replicate the discrete pricing pattern that is observed around sales, which has proved to be a challenge to most price setting models. Simultaneously, the model's learning dynamics are able to reconcile individual price flexibility with an aggregate price level that moves more gradually in response to shocks.

JEL-classification: D21, D83, E31

Key words: Active learning, Price experimentation, Nominal rigidities, Sales, Discrete pricing

1 Introduction

This paper investigates the extent to which the active learning motive of a monopolistic seller who is uncertain about the slope of his demand curve can reconcile the

*I am especially grateful to Wouter den Haan and Sweder van Wijnbergen for their comments and support. Next to that, I also wish to thank Hans Amman, Martin Ellison, Xavier Gabaix, Liam Graham, Peter Leeflang, Bartosz Maćkowiak, Albert Marcet, Filip Matějka, Giuseppe Moscarini, Kevin Sheedy, Christian Stoltenberg, Volker Wieland, Stephen Wright and seminar participants at the NHH-UiO Workshop for stimulating comments and discussions.

†Department of Economics, University of Amsterdam, Valckenierstraat 65-67, 1018 XE Amsterdam, The Netherlands. E-mail: t.willems@uva.nl. Tel.: 0031-20-5257159. Fax: 0031-20-5254254.

volatile, discrete pattern followed by individual prices, with the sluggishness observed in the aggregate price level.

With the arrival of studies analyzing higher-frequency data on prices at the micro level (such as Bils and Klenow (2004)), the type of questions economists have tried to answer has undergone a remarkable twist: while many early contributions focused at explaining why firms are *so reluctant* to change their prices (see *e.g.* Greenwald and Stiglitz (1989), Balvers and Cosimano (1990) and the fixed-price literature), the micro data have shown us that the real puzzle may rather be why firms choose to reprice their products *so often* (*cf.* Eden and Jaremski (2010)). In particular, Kehoe and Midrigan (2010) and Eichenbaum, Jaimovich and Rebelo (2011) report that posted retail prices tend to change once every two to three weeks.

The variability in individual prices is nicely illustrated by Figure 1, which shows the prices posted by a large US retailer for six different goods at a weekly frequency.

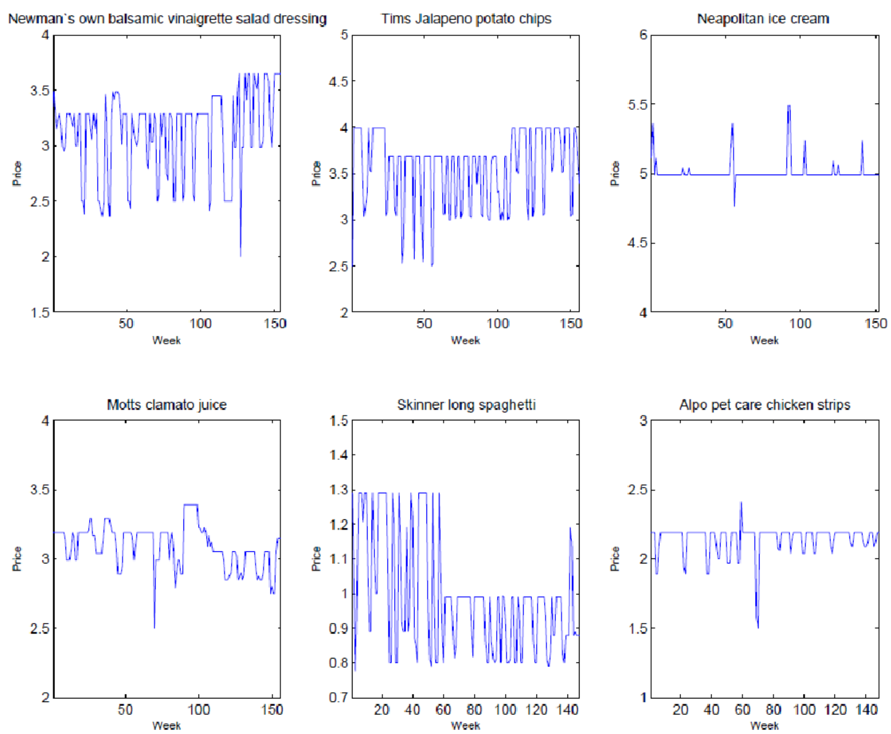

Figure 1: Prices posted by a large US retailer for various goods (source: Eichenbaum, Jaimovich and Rebelo (2008)).

However, next to the aforementioned studies that document the flexibility of individual prices, there is also evidence that the aggregate price level moves more sluggishly in response to shocks (*cf.* Christiano, Eichenbaum and Evans (1999) and Uhlig (2005)). Consequently, it would be nice to have a model that is able to reconcile the rather flexible price path observed at the micro stage, with an aggregate price level that evolves more gradually.

Lately, the profession has made considerable progress in bringing these two (seemingly conflicting) observations together. On the one hand, papers like Gertler and Leahy (2008), Nakamura and Steinsson (2010) and Kehoe and Midrigan (2010) manage to replicate these dimensions of the data by refining the standard menu cost model. On the other hand, another branch of the literature has pointed out that informational imperfections may also play a role in this: Nimark (2008) for example shows that higher order expectations can reconcile individual price flexibility with aggregate price stickiness, while Maćkowiak and Wiederholt (2009) and Matějka (2010a,b) can match these observations by modeling attention as a scarce resource (following Sims (2003)).[1]

However, next to the flexibility of individual prices, higher-frequency price datasets have also shown us that individual price series are characterized by at least one other interesting feature: prices tend to move back and forth between only a few rigid values, thereby showing a lot of discreteness. Matching this dimension of the data with "continuous" models (*i.e.:* models without an exogenously added preference towards discreteness), turns out to be more challenging.[2]

This paper therefore develops a different explanation for the aforementioned micro-macro conflict in pricing behavior - one that is also able to generate individual price series that show a lot of discreteness. It starts by relaxing the rather strong (but standard) assumption that firms can observe their entire demand curve. Instead, in this paper's setting firms can only observe the price-quantity pair they are producing at. Consequently, they are uncertain on the slope of the demand curve for their product. Given this uncertainty, firms are then assumed to be learning about the true value of this parameter - thereby endogenously generating an estimate of their demand curve.

The way in which this learning process is modeled, however differs from the passive approach that Sargent (1999) for example used to analyze the Fed's learning process on the exploitability of the Phillips curve. Under that strategy, the parameter estimation stage is separated from the control stage as a result of which the resulting

[1] Also see Woodford (2009) for a nice bridge between these two strands of the literature.

[2] From the aforementioned papers that manage to reconcile individual price flexibility with aggregate price stickiness, only the ones by Matějka (2010a,b) produce discrete pricing patterns (without an exogenously added preference towards it). Gabaix (2011) also obtains discreteness, but he does not analyze his model's ability to reconcile individual price flexibility with aggregate price stickiness.

learning rule is not optimal.

The present paper on the other hand does optimize the learning rule by taking the links between estimation and control into account. It does so by letting firms learn *actively* about their demand curve. Under active learning, firms realize that they are learning from self-generated observations. Consequently, they take into account that there is a link between the actions they take, and the amount of information that is generated on the parameters they are trying to learn about. As a result, firms become willing to take certain actions that are suboptimal in the short run, because they realize that these actions may generate information - thereby enabling them to achieve higher pay-offs in the future.

It turns out that the active learning motive on the slope of the demand curve gives sellers an incentive to increase the volatility of their control variable (the relative price they post). The reason is that the control variable also serves as the explanatory variable in the regression the sellers are estimating, and the fit of this regression (as for example measured by its R^2) is increasing in the variability of the regressor.

Through this channel, the present paper provides a rationale for the occurrence of sales. This phenomenon has lately received a lot of attention from macroeconomists, as their prevalence has raised the question whether sales are important for the dynamics of the aggregate price level. Other recent papers have explained the existence of sales by referring to the formation of implicit contracts between buyers and sellers (Nakamura and Steinsson, forthcoming), the presence of consumers with different price elasticities (Guimarães and Sheedy, 2011; Chevalier and Kashyap, 2011), or by assuming that price adjustment costs for a sale are lower than those accompanying a regular price change (Kehoe and Midrigan, 2010; Eichenbaum, Jaimovich and Rebelo, 2011). The (complementary) motive proposed by this paper is that firms hold sales in order to generate more information on the slope of their demand curve, so that they can achieve higher profits in the future.

There is indeed evidence that firms are displaying this type of experimental behavior in reality. Pashigian (1988) for example finds that the frequency of sales at US department stores is increasing in the uncertainty on demand, while Campbell and Eden (2010) report that grocers deliberately seem to select extreme prices which they then quickly abandon. This suggests to them that "sellers extensively experiment with their prices" (p.1). This hypothesis is confirmed by survey evidence from the marketing literature: in a questionnaire held among 32 large US retailers, 90 percent of them say to experiment with their price (Gaur and Fisher, 2005). The fact that an increasing number of sellers has also started to employ pricing optimization software (algorithms that tend to take the experimentation motive explicitly into account), only adds to this.[3] More generally, retailers seem to be well aware of the challenges associated with optimal price setting in the face of demand uncertainty (*cf.* Cash and Frankel (1986), Ch. 6) and the revenue management literature has produced a large number of papers that looks at this problem (see Bitran and Caldentey (2003) for an overview and Araman and Caldentey (2009) for a recent example).

Turning to the results of the present paper, the price paths resulting from the experimentation motive show a lot of discreteness: despite the fact that the model consists of continuous equations, prices often move back and forth between only a few rigid values, just as observed in the data (*cf.* Figure 1). The reason is that the price setting rules under active learning take the form of a step function: there is a low price platform (associated with low demand) and a high one (for high demand), separated by a sharp kink in the middle (which serves to avoid the posting of uninformative "moderate" prices). In simulations, the resulting price paths look very much like those generated by Matějka's (2010a,b) rationally inattentive seller and consumer models, as they also manage to produce discrete pricing patterns without having an exogenous preference towards it.[4]

In addition, the step-shaped price setting rules also imply an endogenous form of "local" price stickiness: as long as the model dynamics do not push the seller from the high to the low platform of the policy function (or vice versa), he will keep his price rather constant. Just as in Ss-type models, this implies that prices are rigid in response to small shocks, while they have no difficulties in adjusting to shocks that are large enough to move them into a different region of the policy function.

The remainder of this paper is structured as follows. First, Section 2 will provide a short and non-exhaustive overview on the use of active learning methods in economics (see Kendrick, Amman and Tucci (2011) for a more comprehensive survey) and tries to place the current paper within that literature. Section 3 then describes the model, after which Section 4 explains how the model can be solved. Subsequently, Section 5 contains the calibration and model results, after which a discussion (Section 6) and conclusion (Section 7) complete the paper.

[3] According to the 2004 "Retail Horizons"-survey conducted among 300 US retailers, 36 percent of the responding retailers used pricing optimization software (notable users include Walmart and 7-Eleven, as well as all major hotel and airline companies). In their 2002 survey, this number was only 12 percent. If this trend (along with the increasing importance of online selling) continues, the active learning process of sellers is likely to become even more important in the future as both of these developments greatly facilitate the practical implementation of optimal price experimentation.

[4] As I will explain later on, the discreteness that arises in the present paper is not exact, but looks very much like it. In the rational inattention papers of Matějka (2010a,b) on the other hand, the discreteness is exact. See Matějka and Sims (2010) for a general discussion of discreteness in models with rationally inattentive agents.

2 Relation to the literature

Active learning was first introduced in the economics literature by MacRae (1972) and Prescott (1972).[5] They both considered general optimal control problems in the presence of an experimentation motive. Subsequently, active learning has for example been applied to experimental drug consumption (*cf.* Grossman, Kihlstrom, and Mirman (1977)), investment and growth under uncertainty (Bertocchi and Spagat, 1998) and to optimal monetary policy (see Wieland (2000b), Ellison and Valla (2001) and Svensson and Williams (2007)).

The first application of the active learning-concept to a monopolist who tries to learn his demand curve can be found in Rothschild (1974). He framed the learning process as a two-armed bandit problem and noted that this problem contains a trade-off between long and short run profit maximization (the former calling for price experimentation, while the latter goal is achieved by simply posting the price that present information indicates is most profitable). Subsequently, several papers have managed to establish important analytical results in this context: Mirman, Samuelson and Urbano (1993) for example construct a two-period model in which they derive the conditions under which monopolists are willing to experiment, Trefler (1993) establishes results that determine the direction of experimentation, while Keller and Rady (1999) cast the problem into a continuous time setting and analyze how variations in the discount rate and demand curve switching intensity affect the experimentation incentives.[6]

Following advances in computational power, the present paper is able to take a more practical approach that visualizes the experimentation motive. In this sense, it also relates to Kiefer (1989), who solved Rothschild's (1974) two-armed bandit version of the problem numerically and analyzed how the properties of the value and policy functions varied with the discount factor. In contrast to the present paper, he however did not analyze what this implies for the paths followed by individual and aggregate prices, nor did he establish the link with sales (a phenomenon that did not seem to be of much interest to macroeconomists at the time).

In line with most of the aforementioned contributions, this paper will assume that firms are using their pricing strategy only to learn about the slope of their demand

[5] The concept of active learning is also referred to as "optimal experimentation" (Wieland (2000a), Cosimano (2008)) or "optimal learning" (Wieland, 2000b). It originated in the engineering literature with the contributions of Feldbaum (1960), where it is known as "dual control".

[6] Contemporaneous work by Bachmann and Moscarini (2011) also analyzes a model in which the seller has an experimentation motive, but they mainly use their model to show that fluctuations in first moments can generate fluctuations in uncertainty. I will return to the relevance of Bachmann and Moscarini's work for the present paper in Section 6.

curve. In sharp contrast, there is assumed to be no learning motive on the intercept term.[7] Balvers and Cosimano (1990), which is another important predecessor to the present paper, consider the other polar case in which firms are only trying to learn the intercept of their demand curve in an active manner. By establishing a set of analytical results, they show that this provides firms with an incentive to mute their price adjustments in response to changes in demand. Under their specification, active learning thus implies *the absence of price changes*. Consequently, their model can replicate the delayed response of the aggregate price level in response to shifts in demand, but is unable to match the recently uncovered observation that individual prices change so frequently. This paper's specification on the other hand, in which firms are actively learning the slope of their demand curve, turns out to be able to reproduce both dimensions of the data (see Section 5).

3 Model

Consider a price-setting seller who operates in a monopolistically competitive market. He sells his product in period t at price p_t and serves whatever quantity is demanded at that price. There are no price adjustment costs, so p_t is set on a period-by-period basis in a fully flexible way. In the production process, the seller faces a marginal production cost that is equal to c. For simplicity, the latter is assumed to be known and constant over time. Indicating period t demand for his product by q_t and using δ to denote the discount factor, his profit maximization problem reads:

$$\max_{p_t} \ \mathbb{E}_0 \left\{ \sum_{t=0}^{\infty} \delta^t \left[p_t - c \right] q_t \right\} \tag{1}$$

Demand for his product follows a linear specification. In particular, we have that:

$$\begin{aligned} q_t &= \alpha_t - \beta_t \left[p_t - \overline{p}_t \right] + \gamma q_{t-1} \\ &\equiv \alpha_t - \beta_t r_t + \gamma q_{t-1} \end{aligned} \tag{2}$$

In this equation, $\overline{p}_t$ represents the aggregate price level. Its value lies beyond the control of each individual firm, but can be observed without error. Furthermore, $r_t \equiv p_t - \overline{p}_t$ is the firm's relative price, α_t is the aggregate demand component

[7] Although computational power has increased over the years, solving the more general problem in which firms are learning on both slope and intercept remains very difficult. See Section 6 for a further discussion of this issue and an idea of how the addition of a learning motive for the intercept would affect the results.

that is common to all firms, and β_t is the slope of the demand curve for the good under consideration.[8] Following evidence from the marketing literature (*cf.* Gorden, Goldfarb and Li (2011) and the references therein), this parameter will be time-varying.

There are numerous reasons why the sensitivity of consumers to a certain product's price can vary over time: it may be due to aggregate conditions (consumers becoming more price sensitive in recessions), firm/brand-specific ones (an advertizing campaign making consumers less price sensitive or a competitor opening a nearby store, thereby increasing price competition), while weather conditions may also play a role (think of ice cream when it is hot). In this paper, all of these factors are captured by the stochastic process for β_t (to be specified below).

As both the industrial organization and marketing literature have generated substantial evidence that consumer behavior is subject to some form of brand loyalty (*cf.* Klemperer (1995), Chintagunta, Kyriazidou and Perktold (2001) and the references therein), I allow for the formation of good-specific habits via the parameter γ.[9] For simplicity, the value of this parameter is assumed to be constant and known to the seller. It should however be stressed that the experimentation motive would still be present without any form of habit formation, so from that point of view its inclusion is not essential (see *e.g.* Kiefer (1989), Wieland (2000a) and Bachmann and Moscarini (2011) who all analyze active learning problems without a lagged dependent variable).

The crucial difference with the standard model is the seller's information set. Whereas the standard model makes the rather strong assumption that sellers can perfectly observe the entire demand curve they are facing, this paper relaxes this assumption and keeps the intercept α_t and slope β_t unobserved to the seller. Apart from the (p, q)-point he is producing at every period, the seller thus faces uncertainty on the demand curve for his product. Instead, the seller only knows the stochastic processes that α_t and β_t are assumed to follow.

Although in practice both the intercept and slope of the demand curve are estimated to be persistent (see *e.g.* Smets and Wouters (2007)), this paper makes the simplifying (but counterfactual) assumption that aggregate demand α_t follows an i.i.d.-process with known mean $\overline{\alpha}$, hence:

$$\alpha_t = \overline{\alpha} + \mu_t, \text{ with } \mu_t \sim N(0, \sigma_\mu^2) \tag{3}$$

[8] Formally, β should thus get a j-index (to indicate that it is firm j-specific), but to lighten notation I have omitted the latter here.

[9] To keep the analysis clear, I abstract from the time-consistency issues that are central to Nakamura and Steinsson (forthcoming). As a result, the commitment-channel they emphasize is not at work in the present paper.

This ensures that there is no learning motive for this intercept term (since it has no persistence its expected value is always equal to $\overline{\alpha}$).[10] The most important reason to make this assumption is the fact that it keeps the number of state variables to a minimum - thereby easing the process of finding a numerical solution to the model (see Section 6 for a more detailed discussion of this issue).

Of course, one could establish the exact same by making the opposite counterfactual assumption, namely that α_t is persistent and that β_t follows an i.i.d. process. I find the latter case however less attractive for two reasons. Firstly, aggregate conditions can also be learned relatively easily from other sources such as published statistics, newspaper articles, etc etc. The value of the slope parameter on the other hand is much harder to infer, and apart from the price-quantity observations that firms themselves generate, they have not got much information at their disposal that is useful in learning its true value. It therefore seems natural to assume that firms use their pricing strategy - the only mechanism they have to obtain more information about the slope of their demand curve - for exactly this purpose. Secondly and quite importantly, both the micro data (Pashigian, 1988; Campbell and Eden, 2010) as well as the survey evidence (Gaur and Fisher, 2005) suggest that the active learning process primarily takes place with respect to the slope term.

For these reasons, the unobserved process for β_t is assumed to be persistent. Consequently, sellers in the model do have an incentive to learn more about the actual value of this parameter. I assume that β_t follows an $AR(1)$ with mean $\overline{\beta}$:

$$\beta_t = \left(1 - \rho_\beta\right) \overline{\beta} + \rho_\beta \beta_{t-1} + \eta_t, \text{ with } \eta_t \sim N(0, \sigma_\eta^2) \tag{4}$$

The timing and informational structure of the model are as follows: at the beginning of period t, firms have prior beliefs regarding the unobserved parameter they are trying to learn about, β_t. These prior beliefs can be described by a conditional normal distribution:

$$p\left(\beta_t | \Xi_{t-1}\right) = \mathcal{N}\left(b_{t|t-1}, \Sigma^b_{t|t-1}\right)$$

Here, notation is such that $\mathbb{E}_{t-1}\left[\beta_t\right] \equiv b_{t|t-1}$ and $var_{t|t-1}\left[\beta_t\right] \equiv \Sigma^b_{t|t-1}$. From this expression one can see that Ξ_{t-1}, the vector of state variables describing beliefs, contains two elements: $b_{t|t-1}$ and $\Sigma^b_{t|t-1}$.

Given these beliefs, the firm sets its relative price r_t, after which all shocks materialize and the firm gets to observe realized demand q_t. With this new information,

[10] The i.i.d. assumption for aggregate economic conditions is clearly unrealistic, but assuming that α_t follows a persistent process that is perfectly observable (as a result of which there still is no motive to learn about this term), has the drawback that it adds an extra state variable to the system. As this increases the computational costs to solve the model due to the dimensionality curse without delivering any additional insights, I chose to proceed with the i.i.d.-assumption.

the firm then updates its beliefs via Bayes' rule, which states that:

$$p\left(\beta_t|\Xi_t\right)=\frac{p\left(q_t|\beta_t,r_t,\Xi_{t-1}\right)p\left(\beta_t|\Xi_{t-1}\right)}{p\left(q_t|r_t,\Xi_{t-1}\right)}$$

This implies that we can write the Bayesian updating equations as (*cf.* Zellner (1971)):

$$b_{t|t} = b_{t|t-1}-\frac{\Sigma^b_{t|t-1}r_t}{r_t^2\Sigma^b_{t|t-1}+\sigma^2_\mu}\left(q_t-\gamma q_{t-1}-\overline{\alpha}+b_{t|t-1}r_t\right) \tag{5}$$

$$\Sigma^b_{t|t} = \Sigma^b_{t|t-1}-\frac{\left(\Sigma^b_{t|t-1}\right)^2r_t^2}{r_t^2\Sigma^b_{t|t-1}+\sigma^2_\mu} \tag{6}$$

Finally, in going from the end of period t to the beginning of period $t+1$, we have that:

$$b_{t+1|t} = \left(1-\rho_\beta\right)\overline{b}+\rho_\beta b_{t|t} \tag{7}$$

$$\Sigma^b_{t+1|t} = \rho^2_\beta\Sigma^b_{t|t}+\sigma^2_\eta, \tag{8}$$

Apart from the fact that these updating equations are non-linear functions of q_t and r_t, one should also note that they establish a link between different periods since actions taken in the past affect (the quality of) future beliefs. This adds a dynamic dimension to the firm's pricing problem.

Furthermore, active learning implies that the firm's speed of learning, which is given by the Kalman gain-like expression $\Sigma^b_{t|t-1}r_t/\left(r_t^2\Sigma^b_{t|t-1}+\sigma^2_\mu\right)$ in equation (5), can be affected by the firm's decisions: it is a function of the relative price r_t the firm sets. The reason is that by setting its relative price in an "informative way", a firm can generate more information on the system it is trying to learn about - thereby facilitating the learning process. This contrasts with standard signal extraction problems, such as the one in Lucas (1973), in which the speed of learning is simply a function of the exogenously given signal-to-noise ratio.

In the present setting on the other hand, the signal can be made more informative by introducing additional variability in the control r_t.[11] After all, if one estimates a regression that has the form of equation (2) (which the sellers in this model are

[11] In this respect, active learning also has an interesting relationship with Sims' (2003) idea of rational inattention: while both concepts endogenize the signal-to-noise ratio, rational inattention achieves this by endogenizing the amount of *noise* in the system, while active learning on the other hand endogenizes the *signal*.

doing), one would like the explanatory variable r_t to vary a lot so as to obtain more information on its coefficient β_t.[12] And since the explanatory variable happens to be the control of the system, this can actually be achieved by introducing some additional volatility in this control. This is what is typically referred to as the experimentation effect.

4 Solving the model

The seller's problem is given by (1) subject to equations (2)-(8). As shown by Kiefer and Nyarko (1989), this type of problem still admits a contraction mapping argument. In particular, the problem satisfies Blackwell's sufficiency conditions of monotonicity and discounting, as a result of which it has a fixed point (being the value function). Consequently, it can be solved by dynamic programming. The problem has three state variables ($q_{t-1}, b_{t|t-1}$ and $\Sigma^b_{t|t-1}$), so the Bellman equation can be written as:

$$V\left(q_{t-1}, b_{t|t-1}, \Sigma^b_{t|t-1}\right) = \max_{r_t} \left[\begin{array}{c} \Pi\left(q_{t-1}, b_{t|t-1}, \Sigma^b_{t|t-1}, r_t\right) + \delta \int V\left(q_t, b_{t+1|t}, \Sigma^b_{t+1|t}\right) \\ \times f\left(q_t | q_{t-1}, b_{t|t-1}, \Sigma^b_{t|t-1}, r_t\right) dq \end{array} \right]$$

This equation nicely shows the trade-off between estimation and control that firms face: on the one hand, they want to maximize current period profits Π_t, but on the other hand they also have to take the expected continuation value into account (the second term). The latter contains the potential improvements in future performance that may result from past price experimentation.

The model is solved by iterating over this Bellman equation until the value function converges. The algorithm followed to achieve this is described in Beck and Wieland (2002). It implies that, starting from an initial guess of the value function, one updates the latter by maximizing the right-hand side of the Bellman equation. There, the conditional expectation is evaluated using Gauss-Hermite nodes (since the dependent variable q_t is distributed normally). To speed up convergence, policy iterations are carried out after each value iteration. This reduces the number of value iterations that is needed for convergence, and thereby the number of times that the costly optimization step has to be carried out. Since these type of control problems are typically characterized by multiple local maxima (see Amman and Kendrick

[12] To see this, note from (6) that one would not learn anything about β_t if r_t would remain constant at zero over time. Equation (8) then implies that uncertainty about the true value of β_t will grow with σ^2_η (the variance of the slope innovation) every period. The reason is that setting $r_t = 0$ isolates demand from the slope term β_t as the latter then gets eliminated from the demand curve (2) via the multiplication process.

(1995)), the algorithm starts by conducting a rough grid search, after which a golden section search is carried out to compute the maximum more precisely.

On a 2.79 GHz Pentium processor with 3.46 GB of RAM, the program takes about one hour to converge.

5 Price setting under active learning

We are now in the position to solve the problem set out in Section 3, and see what the introduction of active learning implies for the price setting behavior of firms. In order to do so, Section 5.1 first discusses the model's calibration, after which Section 5.2 describes the model outcomes.

5.1 Calibration

As this paper deals with higher-frequency price movements, I calibrate the model at a weekly frequency. Table 1 summarizes the calibration. The weekly discount factor δ is set to 0.9992 (which implies a standard value for the quarterly discount factor of 0.99). The parameter capturing good-specific habits (γ) is set equal to 0.9875, which is the estimate of Ravn, Schmitt-Grohé and Uribe (2006) converted to a weekly frequency. The marginal cost of production c is normalized to zero.

Symbol	Interpretation	Value
$\overline{\alpha}$	Average aggregate demand	0.38
$\overline{\beta}$	Average slope of demand curve	6
c	Marginal production cost	0
γ	Habit formation coefficient	0.9875
δ	Discount factor	0.9992
$\overline{p}$	Aggregate price level	-1.8
ρ_β	AR-component in β-process	0.99
σ^2_μ	Variance of intercept innovation	0.0164
σ^2_η	Variance of slope innovation	0.0199

Table 1: Baseline calibration.

I will simplify the analysis by taking the aggregate price level $\overline{p}_t$ to be exogenous. In particular, I will just set it equal to a constant, hence $\overline{p}_t = \overline{p} \;\; \forall t$ (the same assumption is effectively made in Matějka (2010a)). One could allow for time variation in the aggregate price level, but this would only obscure the problem without adding much to the analysis.

The actual value of $\overline{p}$, along with the values of the other two parameters that determine the location of the demand curve (the average intercept $\overline{\alpha}$ and average slope $\overline{\beta}$), are however difficult to calibrate as their values cannot be estimated that easily. Consequently, these parameters are set in order to ensure two things: i) that sales q_t do not go negative in simulations and ii) that the average price a firm posts over time approximately equals $\overline{p}$ (so that individual and aggregate prices are internally consistent in the model). I have found values of $\overline{\alpha} = 0.38$, $\overline{\beta} = 6$ and $\overline{p} = -1.8$ to work well for these purposes. These values are not unique and do not have direct economic interpretations (they only determine the location of the demand curve), so the negative value for the aggregate price level has no meaning here.

Whereas the actual value picked for $\overline{\beta}$ is not that important, the relative uncertainty that $\overline{\beta}$ is surrounded with, is key on the other hand as this determines the seller's experimentation incentives. This relative uncertainty can for example be measured by the t-statistic on β_t. The latter has recently been estimated in the marketing literature by Gorden, Goldfarb and Li (2011). Using household panel data running from January 2001 up to December 2006, they come up with an average t-statistic of about 6 across all 19 grocery categories in their study. This implies that β should get a unit standard deviation, which requires σ_η^2 to equal 0.0199.

The autoregressive parameter for β_t is based upon Smets and Wouters (2007), who estimate their equivalent of ρ_β to equal 0.89 on quarterly US data. This leads to a calibrated weekly value of 0.99.

Finally, the variance of aggregate demand α (σ_μ^2) is set equal to the variance of the natural log of linearly detrended US nominal GDP from 1947q1 to 2011q1 (taken from the St. Louis Fed), which equals 0.0164.

5.2 Model results

To see what each separate learning step contributes to the final solution, I will start by reducing the problem described by equations (1)-(8), to two simpler problems.

First, one can derive the policy rule for the standard case in which firms live in a perfect information world where they can observe the true values of α_t and β_t without error. In that case, the problem under consideration reduces to a simpler one and one can show that the profit maximizing price equals:

$$p_t^* = \overline{p} + \frac{1}{2\beta_t}\left[\alpha_t + \beta_t\left(c - \overline{p}\right) + \gamma q_{t-1}\right] - \gamma \mathbb{E}_t\left\{\frac{q_{t+1}}{2\beta_{t+1}}\right\} \tag{9}$$

Here, both the forward looking term ($\gamma \mathbb{E}_t \left\{ q_{t+1}/2\beta_{t+1} \right\}$) as well as the lagged term ($\gamma q_{t-1}/2\beta_t$) result from the dynamic link that is established by the formation of good-specific habits. In particular, this equation shows the trade-off (described in Klemperer (1995)) sellers face between setting a low price to lock consumers into their specific product (which is what the forward looking term captures) and setting a high price to harvest profits by exploiting consumers who are already "addicted" due to past consumption (this is what is expressed by the lagged term).

Next, we can move on to the specification in which α_t and β_t cannot be observed, as a result of which agents have to learn their true values. In that case, firms have the option to follow the so-called "passive" policy rule. It is however key to note that this rule is not optimal as it implies that the seller does not realize that he is learning from self-generated observations. Consequently, he disregards the link between current and future beliefs (*i.e.* in the optimization step he neglects equations (5)-(8)), as a result of which this learning rule lacks any experimentation motives: although firms do update their beliefs as more observations arrive over time, they do not actively seek for better information by posting somewhat more extreme prices.

For this case, the passive price setting rule can be shown to equal:[13]

$$p_t^{\Diamond} = \overline{p} + \frac{1}{2b_{t|t-1}} \left[\overline{\alpha} + b_{t|t-1} \left(c - \overline{p} \right) + \gamma q_{t-1} \right] - \gamma \mathbb{E}_t \left\{ \frac{q_{t+1}}{2b_{t+1|t}} \right\} \tag{10}$$

Deviations of the active learning rule from this passive one, measure the extent of price experimentation.

Unfortunately, it is not possible to obtain an analytical expression for the price setting rule under active learning in the model under consideration. One can however solve for it numerically using the Beck-Wieland algorithm set out in Section 4 of this paper. This was done in order to generate Figure 2. That figure plots the active learning pricing rule and compares it with the passive policy for different combinations of beliefs $\left(b, \Sigma^b\right)$.[14]

The figure clearly shows the experimentation motive: while the passive price setting rule is just a linear, increasing function of demand, the rule resulting from the active learning process contains a kink. This non-linearity shows up in that part of the state space where posting the less extreme, passive learning price would be

[13] Note that the current specification implies that the passive policy rule is equal to the certainty equivalent one (*i.e.* the rule that one obtains if one replaces actual values by estimates in equation (9)). Also see Balvers and Cosimano (1990, p. 887) on this.

[14] Note that the perfect information price setting rule (9) equals the passive learning rule (10) for $b_{t|t-1} = \beta_t$ (this is the certainty equivalence property). Consequently, the passive learning and perfect information rules would coincide in Figure 2 as a result of which I have not included the latter explicitly in the figure.

uninformative on the actual value of the slope term, β_t. Hence, the discontinuity simply serves to avoid uninformative actions (that is: it avoids posting a relative price r_t close to zero).[15] On the other hand, the fact that the policy functions tend to flatten out away from the discontinuity, shows that there is less price experimentation when demand is either relatively high or low. The reason is that in these extreme parts of the state space the relative price picked by the passive policy (which is the optimal one if one neglects the learning motive) has already moved away from zero, due to the fact that demand is so high/low. Consequently, posting the passive price is not that bad anymore from a learning perspective, thereby reducing the need for additional price experimentation (which is costly as it lowers current period profits).

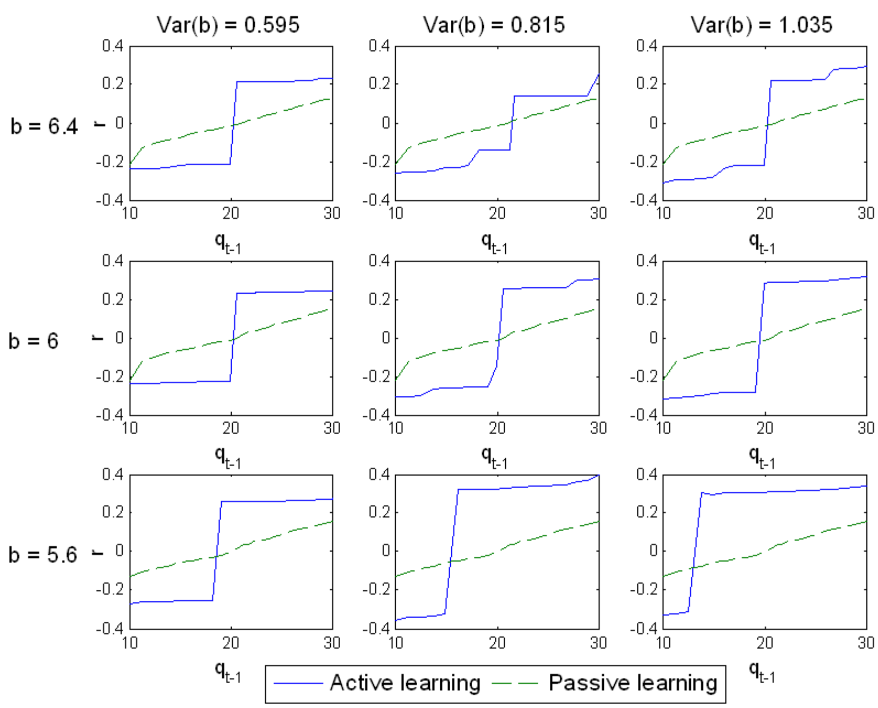

Figure 2: Comparison of active and passive price setting rules for different beliefs.

As a result of this, the optimal price setting rule is shaped like a step function,

[15] More technically, the reason lies in the fact that the value function is twin-peaked. This is due to the fact that it is a combination of a concave function in r_t (static profits) and a convex one (the continuation value). The continuation value is increasing in the absolute value of r_t as more extreme relative prices are more informative. This generates two local maxima (one where $r_t < 0$ and one where $r_t > 0$), from which the max-operator in the Bellman equation then selects the one yielding the highest value (which leads to an avoidance of the region where the continuation value is minimized, *i.e.* that region where r_t is "too close" to zero).

and given beliefs prices tend to jump between two rather rigid values: when demand is low, firms post a low price, while an increase in demand above a certain threshold makes them post the higher price. This can be seen as an endogenous form of "local" price stickiness: as long as the model dynamics do not push a seller from the high to the low price platform of his policy function (or vice versa), he is relatively unresponsive to shocks.

Consequently, the experimentation motive is able to replicate the empirical observation (displayed in Figure 1) that individual prices tend to take on only a couple of values and form a rather discrete pattern over time.[16] This is illustrated in the upper panel of Figure 3, which contains a simulated series (spanning a period of approximately two years) for the control variable r_t under active learning.[17]

Observe that prices tend to change about once every two weeks, which is in line with the findings of Eichenbaum, Jaimovich and Rebelo (2011). By adding a menu cost, one could increase the inertia observed in the price series as these costs would also make experimentation costly in terms of adjustment costs (and not only in terms of forgone current period profits, as is currently the case).

Also note that the experimentation motive is able to generate price series that are more volatile than the underlying marginal costs (which are constant in the present setting). As reported by Eichenbaum, Jaimovich and Rebelo (2011) prices do seem to be more volatile than costs in the data, but standard models have a hard time replicating this feature.

The price path that arises under active learning is quite similar to the one obtained by Matějka (2010a). He assumes that sellers have limited information capacity, as a result of which it becomes optimal for them to pay attention to a source of information that provides a small number of different signals only. Such a signal could be the first digit of unit input costs and since a digit can only take a few different values, there are only a few different signals in his framework - leading to the discreteness.[18]

In the present paper, the (near-)discreteness arises purely from the active learning process: apart from the latter, the model is just a rather standard profit maximization

[16] Since the optimal price setting rule varies a bit with beliefs $\left(b_{t|t-1}, \Sigma^b_{t|t-1}\right)$ and since the upper and lower platforms of these rules are not entirely flat, the resulting pricing pattern is not exactly discrete, but it looks very much like it. By adding a small menu cost one could increase the degree of discreteness even further as the small price changes that can be observed in the current simulation would then no longer be worthwhile (a "region of inaction" would arise).

[17] This particular figure was constructed by giving the seller correct initial beliefs about the true value of β_1, but with an initial uncertainty $\Sigma^b_{1|0} = 0.595$. The typical pattern arising in simulations is robust to choosing different initial beliefs.

[18] In Matějka (2010b), the discreteness follows in a similar way from modeling the consumer as being rationally inattentive.

problem without price adjustment costs. This can also be seen by looking at the two lower panels of Figure 3. Those show the price paths in response to the same series of shocks under both passive learning as well as under perfect information (the case in which sellers can simply observe the true values of α_t and β_t). As one can see, these series no longer show any discreteness (a result of the fact that the passive and perfect information policy rules are not kinked). Since beliefs tend to change more gradually than actual values, the price path in the passive learning specification is basically a smoothed version of the one arising under perfect information. Also note that passively learning sellers have no problems with posting a relative price that is close to zero, as they do not realize that this obstructs their learning process on β_t. As the upper panel of Figure 3 shows, sellers who are learning in the optimal active manner do realize this, as a result of which they avoid posting a zero relative price (this is what the kink in the policy function takes care of).

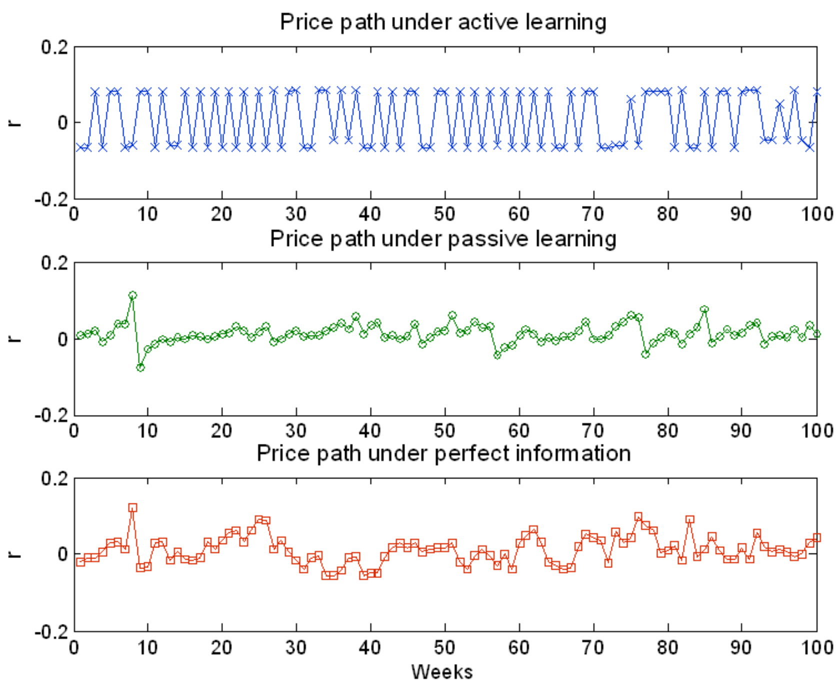

Figure 3: Simulated paths for the relative price posted under active learning, passive learning and under perfect information.

Moreover, since there are no price adjustment costs, prices under passive learning and perfect information change every week, which is too much. With active learning on the other hand, especially the high price tends to show some persistence.[19] This is interesting as the model does not contain any exogenously imposed form of price stickiness. Instead, the kinked nature of the policy function introduces an endogenous form of "local" price stickiness (recall the discussion accompanying Figure 2).

Finally, one can also analyze how the aggregate price level in the active learning model responds to aggregate shocks, and how this response compares to those obtained under passive learning and under perfect information. This is done in Figure 4.[20] This figure shows the posted price averaged over 5,000 sellers. During the simulation, all individual sellers were hit by two types of disturbances: idiosyncratic shocks (all sellers have their individual series for β_t, for example representing the popularity of the particular brand they are selling) and aggregate ones (all these individual series for β_t evolve around a common mean slope of $\overline{\beta}$, capturing aggregate economic conditions). For the first 25 weeks, the common mean value of β_t is set equal to 7. Then, in the 25th week, an aggregate shock suddenly makes all consumers less price sensitive. In particular, $\overline{\beta}$ is assumed to fall from 7 to 5, thereby increasing market power for all sellers.[21]

Several things are to be noted from the figure. First observe that although prices posted by an individual firm show a lot of variation over time due to the experimentation motive (recall Figure 3), the aggregate price level evolves smoothly in the absence of aggregate shocks. The reason is that all experimentation is idiosyncratic to each individual firm, as a result of which it cancels out in the aggregate. Although this probably is an exaggeration of the amount of idiosyncrasy observed in reality, the dataset analyzed in Nakamura (2008) does hint in this direction, as she reports that most of the observed price variation arises at the retail chain level and that only 16 percent of the price variation is common to all stores selling an identical product.

However, when the aggregate shock hits in period 25, the price level does respond in all three specifications. In particular, it goes up, which reflects the increase in market power for sellers associated with the decrease in the average price sensitivity

[19] The fact that this is the high price (and not the low one) is not a general feature of the model, but is rather specific to the current calibration under which the system spends most of its time to the right of the kink in the policy function. To explain the dominance of the high price observed in reality, complementary sale motives (such as the ones offered by Nakamura and Steinsson (forthcoming), Guimarães and Sheedy (2011) and Matějka (2010a,b)) may play a role.

[20] As with Figure 3, this figure was generated assuming that all sellers have correct initial beliefs about the true value of β_1, but with an initial uncertainty $\Sigma^b_{1|0} = 0.595$.

[21] The size of the imposed shock is unrealistically large (especially given the fact that it is assumed to take place within a week), but considering a more gradual change from a high value of $\overline{\beta}$ to a lower one would only increase the sluggishness in the responses of aggregate prices - thereby strengthening the findings that emerge from Figure 4.

of consumers. Next to that, all three models also produce a hump-shaped response to the shock. This is due to the existence of good-specific habits: consumers are "addicted" to the specific good they have been consuming in the past, as a result of which all sellers temporarily obtain the power to charge prices that lie above their new, long run equilibrium values.

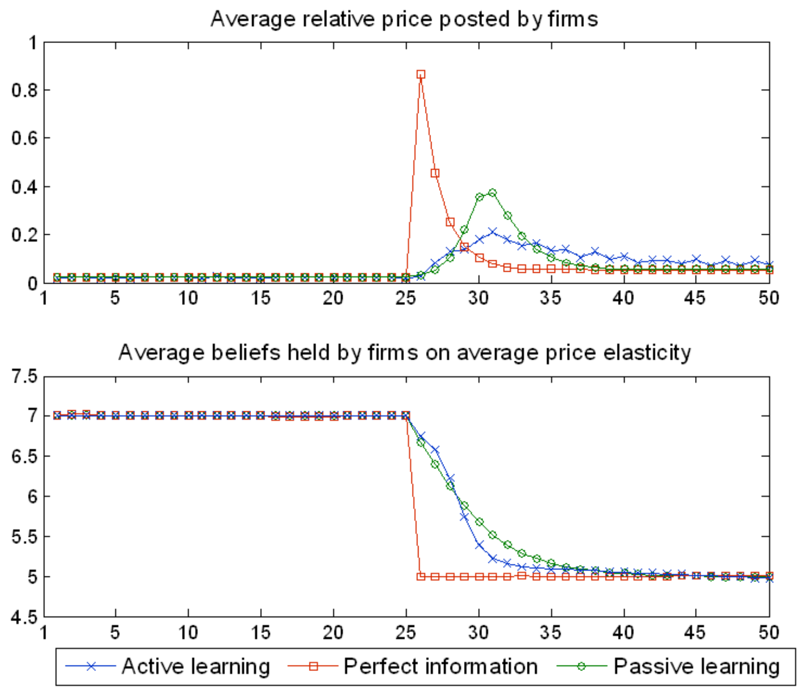

Figure 4: Average posted price (top panel) and average beliefs (bottom panel) under active learning, passive learning and under perfect information.

The persistence of this hump shape however differs strongly between the three models. In particular, the fact that the squared line in the upper panel of Figure 4 converges rather quickly to its new equilibrium value shows that habit formation in itself (that is: without the help of informational imperfections) does not manage to generate that much persistence. As the circled response shows, assuming that the true values of β_t cannot be observed, helps in this respect. The reason is that the learning process implies that firms only gradually find out that the average value of the slope parameter has changed. The circled response was however generated under the assumption that learning takes place in a passive manner. That is: firms post the certainty equivalent price (*i.e.* they simply replace actual parameter values by parameter estimates; compare equations (9) and (10)) and update their beliefs with whatever information this pricing strategy happens to generate.

Under active learning in contrast, firms use their pricing strategy consciously to optimize this learning process. Consequently, learning is faster compared to the passive case (see the lower panel of Figure 4), but average beliefs still need time to converge to the new value of $\overline{\beta}$. As a result, the aggregate price level still moves in an inertial way. Hence, one can conclude that the active leaning process of sellers who are trying to figure out the true value of their slope parameter, is not only able to replicate the volatile, discrete pricing pattern of one individual product (compare the top panel of Figure 3 with Figure 1), but is also consistent with the idea that the aggregate price level responds sluggishly to aggregate shocks.[22]

In this respect, we can thus conclude that the answer to the question "Does flexibility in individual prices translate into aggregate price level flexibility?" should be negative to the extent that individual price changes are driven by the experimentation motive of a seller who is uncertain about the slope of his demand curve.

6 Discussion

Active learning problems can very rapidly become quite challenging to solve. In order to advance, this paper therefore had to abstract from several interesting issues.

First, the analysis in this paper is only partial equilibrium in nature. In this respect, recent work by Bachmann and Moscarini (2011) may form an important contribution. In order to generate "endogenous uncertainty", they manage to embed the active learning problem of a seller in a dynamic stochastic general equilibrium environment. Although this requires them to simplify in some dimensions (they for example reduce the problem to a two-armed bandit setting, where the price elasticity of demand can only take on two values: high or low), the main idea underlying the present paper should carry over to their setup. Investigating whether it is possible to construct a general equilibrium active learning model that is consistent with the behavior of prices at both the micro and macro level, would therefore be a logical next step along this line of research.

Second, this paper also simplifies by assuming that firms are using their pricing strategy to learn about only one of the parameters of the demand curve. This avoids the dimensionality curse that is inherent to the solution method.[23] In particular,

[22] As Figure 4 shows, the latter observation can also be matched by the passive learning model, but recall that that model cannot replicate the volatile, discrete pricing pattern observed at the individual level (see the middle panel of Figure 3), while firms are also not optimizing under that specification (they neglect the production of information).

[23] The problem in which agents are actively learning about two parameters of the demand curve would have six state variables: q_{t-1}, two estimated means, two variances and one covariance esti-

I have assumed that firms are only using their pricing strategy to facilitate the learning process on their slope parameter. As set out in Section 3, I see this as the more realistic problem compared to the other polar case (in which the active learning process only applies to the demand curve's intercept).

Moreover, we already know from Balvers and Cosimano (1990) that active learning on the intercept term provides firms with an incentive to mute their degree of price adjustment in response to changes in demand. Although this enables their model to match the sluggishness observed in the aggregate price level, it is not able to match the volatile behavior of individual prices (nor do the resulting price simulations show any discreteness). The model analyzed in the present paper on the other hand is able to replicate the data along all of these dimensions.

A conceptually similar problem, in which firms are simultaneously learning on both the intercept and slope of an equation that has the form of (2), is studied in Wieland (2006). He analyzes the active learning problem of a central bank who is trying to learn the slope and intercept of the Phillips curve and shows that the extent of experimentation is actually *greater* if learning takes place with respect to both intercept and slope, rather than just with respect to the slope (although the differences are small).[24] This suggests that the results obtained in the present paper are robust to the more general setting in which firms are learning simultaneously on both intercept and slope of the demand curve, although this remains to be verified for the particular problem under consideration of course.

Third, the sale motive proposed by this paper is at best only part of the complete story, as generating information on the demand curve is unlikely to be the only reason for firms to hold sales. It is however important to note that the experimentation motive is fully complementary to other sale motives offered in the literature (such as Nakamura and Steinsson (forthcoming), Guimarães and Sheedy (2011) and Matějka (2010a,b)), so the channel emphasized by this paper is by no means at odds with these earlier contributions. Moreover, as pointed out in footnote 19 of this paper, these complementary channels may be useful in explaining some other dimensions of the data that the experimenting seller model cannot easily account for (such as the dominance of the high price observed in reality).

---

mate. At a weekly frequency (which implies a discount factor close to one), the model then becomes very difficult to solve. Potentially, such a bigger problem could be tackled with faster solution algorithms that rely on approximations to the original problem (such as the ones described in Kendrick (1982), Amman and Kendrick (1994) and Cosimano (2008)).

[24]Because Wieland (2006) does not attempt to match any higher frequency features of the data, he can calibrate his model at a yearly frequency. This allows him to set the discount factor equal to 0.95, which greatly facilitates the convergence of the value function. Nevertheless, he reports that solving his model still takes 3 days on a 2.21 GHz AMD processor with 1.48 GB of RAM.

Fourth, consumers behave rather passively in this paper's model (the consumer side is simply represented by the demand curve (2)). Consequently, consumers do not take into account that sellers have an experimentation motive. If consumers would be aware of this, this could potentially lead to interesting strategic interactions between the two sides of the market - the analysis of which I leave for future work.

Finally, as noted before, the price series produced by this paper's experimenting seller look very similar to the ones produced by Matějka's (2010a,b) rationally inattentive seller and consumer models. Compared to the rational inattention approach, active learning however imposes somewhat more discipline on the model builder: since actively learning sellers are just behaving in a fully rational way, the model does not contain a free parameter whose value is difficult to calibrate (like the information flow constraint in rational inattention models). Nevertheless, the present paper is able to produce rather similar results without such a degree of freedom.

7 Conclusion

This paper has analyzed the optimal pricing strategy of a seller who is actively learning the value of the slope of his demand curve. It is shown that when learning is modeled in an optimal way, the accompanying dynamics are not only useful in matching the behavior of aggregate variables, but also help to replicate the micro-dimensions of the data. In particular, as achieved by Matějka (2010a,b), the model constructed in this paper manages to generate all of the following three stylized facts that seem present in reality:

1. Individual prices change frequently.
2. Individual prices tend to move back and forth between only a few rigid values.
3. The aggregate price level responds gradually to aggregate shocks.

The model can match stylized fact #1 because the introduction of active learning gives firms an incentive to experiment with their price. The reason is that choosing a rather volatile price path facilitates the learning process - just as variability in an explanatory variable facilitates the estimation of a regression equation. Next to the fact that such a form of price experimentation is optimal for sellers wishing to maximize their firm's value, there is also empirical evidence that sellers are displaying this type of behavior in reality (see Pashigian (1988, who documented a positive impact of demand uncertainty on the frequency of sales), Campbell and Eden (2010, who analyzed US scanner data) and Gaur and Fisher (2005, who held a survey among large US retailers)). In line with the dataset analyzed by Nakamura (2008), most of the price variation in the model also arises from the retail level.

In addition, the experimenting seller model can go a long way in reproducing the discreteness of individual prices observed in reality (stylized fact #2). Replicating this feature of the data has proved to be a major challenge to most price setting models. Although the discreteness arising in this paper is not exact, it looks very much like it. The reason is that the experimentation motive shapes the price setting rules like a step function: they have a high price platform (associated with high demand) and a low one (associated with low demand), separated by a sharp kink. Consequently, the model displays a form of endogenous, local price stickiness (despite the absence of price adjustment costs): as long as a shock does not move a seller from the high to the low platform (or vice versa), it is optimal for him to keep his price rather fixed - even if demand varies locally.

Finally, notwithstanding the flexibility of individual prices under active learning, the aggregate price level still reacts sluggishly to shocks. This is due to the fact that average beliefs change only gradually over time. Consequently, the model is also able to capture the inertia found in prices at the aggregate level (stylized fact #3).[25] To the extent that price changes at the individual level are driven by the experimentation motive of a seller who is actively learning the slope of his demand curve, the flexibility of prices at the micro stage thus does not translate into aggregate price level flexibility.

8 References

Amman, H.M. and D.A. Kendrick (1994), "Active Learning: Monte Carlo Results", *Journal of Economic Dynamics and Control*, 18 (1), pp. 119 - 124.

Amman, H.M. and D.A. Kendrick (1995), "Nonconvexities in Stochastic Control Models", *International Economic Review*, 36 (2), pp. 455-475.

Araman, V.F. and R. Caldentey (2009), "Dynamic Pricing for Nonperishable Products with Demand Learning", *Operations Research*, 57 (5), pp. 1169-1188.

Bachmann, R. and G. Moscarini (2011), "Business Cycles and Endogenous Uncertainty", mimeo, Yale University.

Balvers, R.J. and T.F. Cosimano (1990), "Actively Learning about Demand and the Dynamics of Price Adjustment", *Economic Journal*, 100 (402), pp. 882-898.

[25]In conversations, I always liked to simplify the message of this paper as follows: this paper presents a model of "active learning". The "active"-part of the setup delivers the volatility and (near-)discreteness of individual prices, while the fact that there is still some "learning" going on, generates the sluggish response of the price level to aggregate shocks.

Beck, G.W. and V. Wieland (2002), "Learning and Control in a Changing Economic Environment", *Journal of Economic Dynamics and Control*, 26 (9-10), pp. 1359-1377.

Bertocchi, G. and M. Spagat (1998), "Growth Under Uncertainty With Experimentation", *Journal of Economic Dynamics and Control*, 23 (2), pp. 209–231.

Bils, M. and P.J. Klenow (2004), "Some Evidence on the Importance of Sticky Prices", *Journal of Political Economy*, 112 (5), pp. 947-985.

Bitran, G. and R. Caldentey (2003), "An Overview of Pricing Models for Revenue Management", *Manufacturing and Service Operations Management*, 5 (3), pp. 203-229.

Campbell, J.R. and B. Eden (2010), "Rigid Prices: Evidence from US Scanner Data", Federal Reserve Bank of Chicago Working Paper No. 2005-08.

Cash, R.P. and H.H. Frankel (1986), *Improving Apparel Shop Profits*, National Retail Merchants Association.

Chevalier, J.A. and A.K. Kashyap (2011), "Best Prices", Chicago Booth Working Paper No. 11-01.

Christiano, L.J., M. Eichenbaum and C.L. Evans (1999), "Monetary Policy Shocks: What Have We Learned and to What End?", in: J. Taylor and M. Woodford (*eds.*), *Handbook of Macroeconomics*, Vol. 1A, Amsterdam, North-Holland: Elsevier, pp. 65-148.

Cosimano, T.F. (2008), "Optimal Experimentation and the Perturbation Method in the Neighborhood of the Augmented Linear Regulator Problem", *Journal of Economic Dynamics and Control*, 32 (6), pp. 1857-1894.

Eden, B. and M.S. Jaremski (2010), "Discreteness and Nominal Rigidity: Do Supermarket Prices Move Too Much?", mimeo, Vanderbilt University.

Eichenbaum, M., N. Jaimovich and S. Rebelo (2008), "Reference Prices and Nominal Rigidities", CEPR Working Paper No. 6709.

Eichenbaum, M., N. Jaimovich and S. Rebelo (2011), "Reference Prices and Nominal Rigidities", *American Economic Review*, 101 (1), pp. 234-262.

Ellison, M. and N. Valla (2001), "Learning, Uncertainty and Central Bank Activism in an Economy with Strategic Interactions", *Journal of Monetary Economics*, 48 (1), pp. 153-171.

Feldbaum, A.A. (1960), "Dual Control Theory", *Automation Remote Control*, 21, pp. 874-880.

Gabaix, X. (2011), "A Sparsity-Based Model of Bounded Rationality", mimeo, NYU.

Gaur, V. and M.L. Fisher (2005), "In-Store Experiments to Determine the Impact of Price on Sales", *Production and Operations Management*, 14 (4), pp. 377-387.

Gertler, M. and J. Leahy (2008), "A Phillips Curve with an Ss Foundation", *Journal of Political Economy*, 116 (3), pp. 533-572.

Gorden, B., A. Goldfarb and Y. Li (2011), "Does Price Elasticity Vary with Economic Growth? A Cross-Category Analysis", mimeo, Columbia University.

Greenwald, B. and J.E. Stiglitz (1989), "Toward a Theory of Rigidities", *American Economic Review*, 79 (2), pp. 364-369.

Grossman, S., R. Kihlstrom and L. Mirman (1977), "A Bayesian Approach to the Production of Information and Learning By Doing", *Review of Economic Studies*, 44 (3), pp. 533–547.

Guimarães, B. and K. Sheedy (2011), "Sales and Monetary Policy", *American Economic Review*, 101 (2), pp. 844-876.

Kehoe, P.J. and V. Midrigan (2010), "Prices Are Sticky After All", Research Department Staff Report 413, Federal Reserve Bank of Minneapolis.

Keller, G. and S. Rady (1999), "Optimal Experimentation in a Changing Environment", *Review of Economic Studies*, 66 (3), pp. 475-507.

Kendrick, D. A. (1982), "Caution and Probing in a Macroeconomic Model", *Journal of Economic Dynamics and Control*, 4, pp. 149-170.

Kendrick, D.A., H.M. Amman and M.P. Tucci (2011), "Learning about Learning in Dynamic Economic Models", *Handbook of Computational Economics*, forthcoming.

Kiefer, N.M. (1989), "A Value Function Arising in the Economics of Information", *Journal of Economic Dynamics and Control*, 13 (2), pp. 201-223.

Kiefer, N.M. and Y. Nyarko (1989), "Optimal Control of an Unknown Linear Process with Learning", *International Economic Review*, 30 (3), pp. 571-586.

Klemperer, P. (1995), "Competition when Consumers have Switching Costs: An Overview with Applications to Industrial Organization, Macroeconomics, and International Trade", *Review of Economic Studies*, 62 (4), pp. 515-539.

Lucas, R.E. (1973), "Some International Evidence on Output-Inflation Tradeoffs", *American Economic Review*, 63 (3), pp. 326-334

Maćkowiak, B. and M. Wiederholt (2009), "Optimal Sticky Prices under Rational Inattention", *American Economic Review*, 99 (3), pp. 769-803.

MacRae, E.C. (1972), "Linear Decision With Experimentation", *Annals of Economic and Social Measurement*, 1, 437–448.

Matějka, F. (2010a), "Rationally Inattentive Seller: Sales and Discrete Pricing", mimeo, CERGE-EI.

Matějka, F. (2010b), "Rigid Pricing and Rationally Inattentive Consumer", mimeo, CERGE-EI.

Matějka, F. and C.A. Sims (2010), "Discrete Actions in Information-Constrained Tracking Problems", mimeo, Princeton University.

Mirman, L.J., L. Samuelson and A. Urbano (1993), "Monopoly Experimentation", *International Economic Review*, 34 (3), pp. 549-563.

Nakamura, E. (2008), "Pass-Through in Retail and Wholesale", *American Economic Review: Papers and Proceedings*, 98 (2), pp. 430-437.

Nakamura, E. and J. Steinsson (2010), "Monetary Non-Neutrality in a Multisector Menu Cost Model", *Quarterly Journal of Economics*, 125 (3), pp. 961-1013.

Nakamura, E. and J. Steinsson (forthcoming), "Price Setting in Forward-Looking Customer Markets", *Journal of Monetary Economics*.

Nimark, K. (2008), "Dynamic Pricing and Imperfect Common Knowledge", *Journal of Monetary Economics*, 55 (2), pp. 365-382.

Pashigian, B.P. (1988), "Demand Uncertainty and Sales: A Study of Fashion and Markdown Pricing", *American Economic Review*, 78 (5), pp. 936-953.

Prescott, E.C. (1972), "The Multi-Period Control Problem under Uncertainty", *Econometrica*, 40 (6), pp. 1043–1058.

Ravn, M.O., S. Schmitt-Grohe and M. Uribe (2006), "Deep Habits", *Review of Economic Studies*, 73 (1), pp. 195-218.

Rothschild, M. (1974), "A Two-Armed Bandit Theory of Market Pricing", *Journal of Economic Theory*, 9 (2), pp. 185-202.

Sargent, T.J. (1999), *The Conquest of American Inflation*, Princeton and Oxford: Princeton University Press.

Sims, C.A. (2003), "Implications of Rational Inattention", *Journal of Monetary Economics*, 50 (3), pp. 665-690.

Smets, F. and R. Wouters (2007), "Shocks and Frictions in US Business Cycles: A Bayesian DSGE Approach", *American Economic Review*, 97 (3), pp. 586-606.

Svensson, L.E.O. and N. Williams (2007), "Bayesian and Adaptive Optimal Policy under Model Uncertainty", mimeo, Princeton University.

Trefler, D. (1993), "The Ignorant Monopolist: Optimal Learning with Endogenous Information", *International Economic Review*, 34 (3), pp. 565-581.

Uhlig, H. (2005), "What Are the Effects of Monetary Policy on Output? Evidence from an Agnostic Identification Procedure", *Journal of Monetary Economics*, 52 (2), pp. 381-419.

Wieland, V. (2000a), "Learning by Doing and the Value of Optimal Experimentation", *Journal of Economic Dynamics and Control*, 24 (4), pp. 501-534.

Wieland, V. (2000b), "Monetary Policy, Parameter Uncertainty and Optimal Learning", *Journal of Monetary Economics*, 46 (1), pp. 199-228.

Wieland, V. (2006), "Monetary Policy under Uncertainty about the Natural Unemployment Rate: Brainard-Style Conservatism versus Experimental Activism", *Advances in Macroeconomics*, 6 (1), Article 1.

Woodford, M. (2009), "Information-Constrained State-Dependent Pricing", *Journal of Monetary Economics*, 56 (S1), pp. 100-124.

Zellner, A. (1971), *Introduction to Bayesian Inference in Econometrics*, Wiley.