Regular Prices and Sales*

Paul Heidhues  
ESMT and CEPR

Botond Kőszegi  
University of California, Berkeley

September 17, 2010

Abstract

We study the properties of a profit-maximizing monopolist’s optimal price distribution when selling to a loss-averse consumer, where (following Kőszegi and Rabin (2006)) we assume that the consumer’s reference point is her recent rational expectations about the purchase. If it is close to costless for the consumer to observe the realized price of the product, the monopolist chooses low and variable “sale” prices with some probability and a high and sticky “regular” price with the complementary probability—a pattern that is consistent with several recently documented facts regarding supermarket pricing. Realizing that she will buy at the sale prices and hence that she will purchase with positive probability, the consumer chooses to avoid the painful uncertainty in whether she will get the product by buying also at the regular price. If it is more costly for the consumer to observe the realized price, a sale is less tempting and hence less effective in generating an expectation to buy with positive probability, so that the monopolist chooses a sticky price and holds no sales—a pattern that is consistent with the pricing behavior of some other retailers (e.g. movie theaters). We also show that ex-ante competition for loyal consumers leads to sticky pricing while ex-post competition leads to marginal-cost pricing, and discuss several other extensions of the model.

Keywords: Reference-dependent utility, gain-loss utility, loss aversion, sticky prices, sales, supermarket pricing.

---

*We thank Stefano DellaVigna, Erik Eyster, Yoram Halevy, Ulrich Kamecke, Emir Kamenica, Matthew Rabin, Andrei Shleifer, and Joel Sobel for enlightening discussions, and seminar audiences at Arizona, Central European University, Berkeley, ITAM, LSE, Munich, Pompeu Fabra, UCLA, and UCSD, as well as conference participants at the Bounded Rationality in Economics Conference in Austin, Conference in Behavioral Economics in Jerusalem, and Theoretischer Ausschuss des VfS, for comments. Heidhues gratefully acknowledges Berkeley’s hospitality while this paper was written, and financial support from the Deutsche Forschungsgemeinschaft through SFB/TR-15. Kőszegi thanks the Central Bank of Hungary for hospitality while some of this work was completed, and the National Science Foundation for financial support under Award #0648659.
1 Introduction

Consumer retail prices exhibit a systematic combination of stickiness and flexibility. The stickiness of prices—that they can remain unchanged for many months despite changing cost or demand circumstances—has long been considered a stylized fact in industrial organization and macroeconomics. Yet more recent research on supermarket pricing has qualified this stylized fact by documenting that although regular prices do change rarely, there are also frequent sales, and sale prices are much more variable than regular prices. At the same time, this qualification does not seem to apply to many other types of retailers—such as movie theaters or restaurants—which simply have sticky prices and few non-cyclical sales.\footnote{We review evidence for these claims in Section 2}

While existing theories have clearly identified many important considerations in retailer pricing decisions, to our knowledge there is no micro-founded model that convincingly explains in one framework the above puzzling combination of stickiness and flexibility. In this paper, we develop a potential explanation for all the above patterns by introducing consumer loss aversion into a simple classical environment of monopolistic pricing. We assume that a risk-neutral profit-maximizing monopolist sells a single product to a representative consumer with known valuation, and the consumer’s reference point for evaluating her purchase is her recent rational expectations about the purchase. If the consumer automatically finds out the product’s price—such as at a supermarket she visits whether or not she buys this particular product—the monopolist charges low sale prices with some probability and a high regular price with the complementary probability. The sale prices are chosen such that it is not credible for the consumer not to buy at these prices. Then, because the consumer expects to purchase with positive probability and dislikes uncertainty in whether she will get the product, she chooses to buy also at the regular price. In contrast, if the consumer does not automatically know whether the product is on sale—such as for a movie theater she visits only if she is going to watch a movie—a sale is less tempting and hence less effective in making not buying non-credible for her, so that the monopolist simply sets a regular price and has no sales. In either case, because the consumer dislikes uncertainty in how much she pays, to get her to choose to buy at the regular price the monopolist makes the regular price sticky.
After reviewing the key empirical evidence on pricing in Section 2, in Section 3 we present our basic model of supermarket pricing, which uses the framework of Köszegi and Rabin (2006) to incorporate consumer loss aversion into a simple model of first-degree price discrimination. There is a single product and a single representative consumer. If the consumer gets the product, she derives consumption utility $v$ from it, and she also derives additive consumption disutility from any money she pays. In addition, the consumer derives gain-loss utility from the comparison of her consumption utility in the product and money dimensions to a reference point equal to her lagged expectations regarding the same outcomes, with losses being more painful than equal-sized gains are pleasant. Suppose, for example, that the consumer had been expecting to buy the product for either $5 or $7. If she buys it for $6, she experiences no gain or loss in the product dimension and “mixed feelings” in the money dimension consisting of a loss relative to the possibility of paying $5 and a gain relative to the possibility of paying $7, with the weight on the loss equal to the probability with which she had been expecting to pay $5. If she does not buy, she experiences a loss in the product dimension and (paying $0) a gain in the money dimension relative to both prices $5 and $7. To determine expectations and behavior with these preferences, we assume that the consumer must form credible purchase plans: given the expectations induced by her plan of which prices to buy at, buying at exactly those prices must be optimal. Accordingly, we assume the consumer chooses a credible plan—called a preferred personal equilibrium or PPE—that maximizes her ex-ante expected utility.

The above consumer interacts with a risk-neutral profit-maximizing monopolist with deterministic production cost. In period 0, the monopolist commits to a price distribution. This commitment assumption captures, in a reduced form, the idea that a patient firm would have the incentive to develop a reputation for playing the long-run optimal price distribution. The consumer observes the price distribution while forming expectations about her own price-contingent behavior. In period 1, a price is drawn from the distribution, and the consumer decides whether to buy a single item of the good. For technical reasons, we assume that the price distribution must be discrete with atoms at least $\Delta > 0$ apart, and look for the limit-optimal price distribution as $\Delta$ approaches zero.

We analyze our basic model in Section 4. In Section 4.1 we show that for any loss-averse
preferences by the consumer, the monopolist’s limit-optimal price distribution consists of a region of continuously distributed low sale prices and a single (atomic) high regular price (see Figure 1). We explain the intuition in three parts.

First, despite a loss-averse consumer’s dislike of uncertainty—in fact, by exploiting this dislike—the monopolist can earn greater profits by charging uncertain prices than by charging a deterministic price. If the monopolist uses a deterministic price $p$, then it cannot earn revenue of more than $v^2$. But consider instead the strategy of sometimes charging sale prices low enough to make not buying

---

2 In this case, any rational expectations match actual behavior, so in PPE gain-loss utility must be zero. As a result, the consumer prefers to maximize consumption utility, not buying if $p > v$. And such a plan is credible: once the consumer makes her preferred plan not to buy, she would experience paying for the product as a painful loss, so that she would especially not like to buy.
at these prices non-credible, and at other times charging a high regular price. Realizing that she
will buy at the sale prices, would the consumer also buy at the regular price? With a plan not to
do so, she expects to get the product with an interior probability, so she feels a pleasant gain if she
gets it and an unpleasant loss if she does not get it. Due to loss aversion, she feels the loss more
heavily, so that she receives negative expected gain-loss utility in the product dimension. Hence,
she prefers to eliminate uncertainty in whether she will get the product, and is therefore willing to
buy at all prices even if the regular price exceeds \( v \) somewhat. More surprisingly, by exploiting a
type of time inconsistency to push the consumer’s expected utility below zero, the firm can lead her
to pay an *average* price exceeding \( v \). When the consumer decides to buy at a sale price in period 1,
she does not take into account that this increases her period-0 expectations to consume and spend
money, lowering her expected utility. In this sense of leading the consumer to choose outcomes she
does not like ex ante, the monopolist’s pricing strategy is manipulative.

Second, the profit-maximizing way to execute the above “luring sales” is to put a small weight
on each of a large number of sale prices. If the consumer had expected not to buy, she would
experience paying for the product as a loss and getting the product merely as a gain, creating a low
willingness to pay for the product. To make not buying non-credible, then, the monopolist puts a
small weight on a low price \( p \) chosen such that even if the consumer expected not to buy, she would
buy at \( p \). Given that the consumer realizes she will buy at \( p \), she experiences not getting the product
partially as a loss rather than a foregone gain, and paying for it partially as a foregone gain rather
than a loss, increasing her willingness to pay. As a result, not buying at a slightly higher price is
also non-credible, allowing the monopolist to charge higher prices at all other times. Continuing
this logic further, the monopolist needs to charge each sale price with only a low probability.

Third, the regular price is sticky because—just as she dislikes uncertainty in whether she gets
the product—the consumer dislikes uncertainty in how much she pays: she experiences a gain if the
regular price is lower than it could have been and a loss if the regular price is higher than it could
have been, and due to loss aversion the latter is more heavily felt. Because this would decrease the
consumer’s willingness to buy at the regular price in addition to the sale prices, it creates a strong
incentive for the firm to eliminate variation in the regular price.
While the bulk of our paper is devoted to explaining the combination of stickiness and flexibility in supermarket prices, our framework also allows us to make predictions on the pricing of different kinds of retailers by considering natural alternatives to our basic assumptions. In Section 4.2 we provide a potential explanation for the difference in the frequency of sales at supermarkets and some other retailers (such as movie theaters) based on the ease with which a consumer observes whether a product is on sale. Once a consumer is at the supermarket to buy her usual supplies, she automatically sees which other items are on sale that day and will be tempted to buy them, so that a sale is an effective way to make a plan of not buying non-credible. But since a moviegoer does not go to the theater other than to watch a movie, before she decides to go she would not as easily know whether a movie is on sale that day, making a sale less tempting. As a result, a sale is less effective in making a plan of not buying non-credible. We formalize this distinction by showing that when the consumer must pay a high-enough monetary or effort cost to learn the price realization in period 1, the monopolist chooses a sticky price and holds no sales.

In Section 5 we explore a number of further extensions and modifications of our model. First, to demonstrate some robustness of our stickiness result in an environment that with classical consumers would generate price variation, we study pricing when consumers have heterogeneous consumption values and the firm has uncertain marginal cost. While we are not able to solve for the fully optimal price distribution, in a simple restricted class of distributions we show that if the firm’s cost is sufficiently narrowly distributed, the firm still chooses a sticky regular price and substantially lower and variable sale prices. In addition, we consider perfect ex-ante competition for consumers: two firms simultaneously announce their price distributions in period 0, and in the same period the consumer decides which firm to visit in period 1. In this case, firms will compete for consumers by offering a sticky price. This prediction may explain, for instance, why restaurants—typically facing strong local competition—have far fewer sales than supermarkets. In contrast, perfect ex-post competition—where consumers buy a cheapest product on the market—generates marginal-cost pricing, potentially explaining, for instance, why economy-class airline tickets have highly variable prices. We also discuss many other variants of our main model, including the possibility that gain-loss utility is lower in money than in the product, that consumers enter the
market with non-rational initial expectations, and that the monopolist faces a competitive fringe.

In Section 6, we summarize the behavioral-economics and pricing literatures most related to our paper. While we believe many other theories capture realistic aspects of firm pricing, we argue that none explain the pattern of regularities our model does, so that consumer loss aversion is also likely part of the story. We conclude the paper in Section 7 by pointing out some pricing patterns that our model cannot explain.

2 Summary of Evidence

In this section, we summarize some of the key evidence on the price patterns of consumer retail products.

2.1 (Regular) Retail Prices are Sticky

The notion that consumer retail prices are sticky—that they change only intermittently despite presumably more rapid changes in the economic environment that one would expect to change demand elasticities and hence to create incentives for changing prices—has long been an accepted stylized fact in industrial organization. In a classic study, Cecchetti (1986) finds that the time between magazine price changes is typically over a year and sometimes over a decade. For a selection of goods in a mail-order catalog, Kashyap (1995) observes an average of 14.7 months between price changes. MacDonald and Aaronson (2006) document that for restaurant prices, the median duration between price changes is around 10 months. Even at the lower end of the stickiness spectrum, Bils and Klenow (2004) find a median price duration of 4.3 months for non-shelter items in the Bureau of Labor Statistics (BLS) data underlying the Consumer Price Index.

In supermarkets, regular prices change about once a year (Kehoe and Midrigan 2008, Eichenbaum, Jaimovich and Rebelo 2008). And as an example of the stickiness of prices in the face of significant changes in circumstances, Eichenbaum et al. (2008) document that conditional on the weekly price being constant and equal to the regular price, the standard deviation of quantities sold is 42%.
2.2 Prices at Supermarkets Feature Frequent Sales with Variable Prices

Although *regular* prices at supermarkets are quite stable, *posted* prices change every two or three weeks on average, typically by moving away from the regular price and then quickly returning to it (Chevalier, Kashyap and Rossi 2003, Kehoe and Midrigan 2008, Eichenbaum et al. 2008). Furthermore, most of these temporary price changes are sales (decreases) rather than increases, with the mean deviation being -22% of the regular price (Kehoe and Midrigan 2008).

Not only are sales frequent, sale prices are less sticky than regular prices. Klenow and Kryvtsov (2008) document that it is more likely for a sale price to change from one promotion to the next than for a regular price to change when interrupted by a sale. Nakamura and Steinsson (2009) find that for the median product category, the sale price changes in 48.7 percent of the weeks during a multi-week sale, while the regular price changes in only 6.1 percent of weeks. Similarly, the number of unique prices as a fraction of total weeks spent on sale is 0.434, while the same number for regular prices is 0.045.

2.3 At Many Retailers, Sales are Less Common than in Supermarkets

The frequency of sales that has been observed at supermarkets does not seem to be a general feature of consumer retail prices—many retailers simply charge a sticky price and rarely have non-cyclical sales. Movies, for instance, largely sell at the same price for extended periods of time (Einav and Orbach 2007). Similarly, many previous studies of price stickiness, including the Cecchetti (1986) study on newspapers and the MacDonald and Aaronson (2006) study on restaurants mentioned above, do not seem to find frequent sales. And while Eichenbaum et al. (2008) report that sale prices constitute about 32% of price observations at supermarkets, Klenow and Kryvtsov (2008) find that overall they constitute only 8% of price observations.

Of course, some retailers have cyclical sales: for instance, movie theaters have matinees and clothes retailers have clearance sales. As we discuss in Section 5.3 because these sales are responses to obvious substantial changes in demand predicted by most consumers, from the perspective of our model they are qualitatively different from sales at supermarkets.
3 Model

In this section, we introduce our basic model of pricing with a loss-averse consumer. A risk-neutral profit-maximizing monopolist is looking to sell a single product with deterministic production cost $c$ to a single representative consumer. In Section 5.1 we allow for consumer heterogeneity and cost uncertainty, and in Section 5.2 we consider competition between sellers. The interaction between the monopolist and the consumer lasts two periods, 0 and 1. In period 0, the monopolist commits to a price distribution $\Pi(\cdot)$ for its product. The consumer learns the price distribution and (in a way detailed below) forms stochastic beliefs regarding her purchase. In period 1, a price $p$ is drawn from $\Pi(\cdot)$, and after observing the price, the consumer decides whether to buy a single item of the product, choosing quantity $b \in \{0, 1\}$.

Our assumption that the firm can commit to the price distribution captures, in a static reduced form, a patient firm’s dynamic incentives to forego possible short-term profits to manage consumers’ price expectations. In Appendix A we provide micro-foundations for this assumption based on Fudenberg and Levine (1989), in which the firm can develop a “reputation” for playing the optimal committed price distribution. More generally, it seems plausible to assume that a patient firm realizes that over time, consumers will learn the distribution of prices and incorporate it into their expectations.

To ensure the existence of an optimal price distribution and to make our proofs technically easier, we make two assumptions that do not seem to affect the forces driving our results. First, we assume that the monopolist must choose a discrete price distribution in which neighboring atoms are at least $\Delta > 0$ apart. Second, we assume that any indifference by the consumer in period 1 is broken in favor of buying. Without these assumptions, the optimal price distribution would not exist, although even in that case price distributions close to what we find would approximate the least upper bound on profits arbitrarily closely. In Appendix B we identify properties of the optimal price distribution for $\Delta > 0$, but in the text we state these results in a more transparent form, in the limit as $\Delta$ approaches zero:

**Definition 1.** The price distribution $\Pi(\cdot)$ is limit-optimal if there exist a sequence $\Delta_i \to 0$ and optimal price distributions $\Pi_i(\cdot)$ for each $\Delta_i$ such that $\Pi_i \to \Pi$ in distribution.
Our model of consumer behavior follows the approaches of K˝ oszegi and Rabin (2006) and Heidhues and K˝ oszegi (2008), but it adapts and simplifies these theories to fit the decision of whether to purchase a single product. The consumer’s utility function has two components. Her consumption utility is \((v - p)b\), so that the consumption value of the product is \(v\). Consumption utility can be thought of as the classical notion of outcome-based utility. In addition, the consumer derives gain-loss utility from the comparison of her period-1 consumption outcomes to a reference point given by her period-0 expectations (probabilistic beliefs) about those outcomes. Let \(k^v = vb\) and \(k^p = -pb\) be the consumption utilities in the product and money dimensions, respectively. For a riskless consumption outcome \((k^v, k^p)\) and riskless expectations \((r^v, r^p)\) defined over the two dimensions of consumption utility, total utility is

\[
u(k^v| r^v) + u(k^p| r^p) = k^v + \mu(k^v - r^v) + k^p + \mu(k^p - r^p).
\] (1)

We assume that \(\mu\) is two-piece linear with a slope of \(\eta > 0\) for gains and a slope of \(\eta\lambda > \eta\) for losses. By positing a constant marginal utility from gains and a constant and larger marginal disutility from losses, this formulation captures prospect theory’s (Kahneman and Tversky 1979, Tversky and Kahneman 1991) loss aversion, but ignores prospect theory’s diminishing sensitivity. The parameter \(\eta\) can be interpreted as the weight attached to gain-loss utility, and \(\lambda\) as the coefficient of loss aversion.

Beyond loss aversion, our specification in Equation 1 incorporates two further assumptions. First, the consumer assesses gains and losses in the two dimensions, the product and money, separately. Hence, if her reference point is not to get the product and not to pay anything, for instance, she evaluates getting the product and paying for it as a gain in the product dimension and a loss in the money dimension—and not as a single gain or loss depending on total consumption utility relative to the reference point. This is consistent with much experimental evidence commonly interpreted in terms of loss aversion.\footnote{Specifically, it is key to explaining the endowment effect—that randomly assigned “owners” of an object value it more highly than “non-owners”—and other observed regularities in trading behavior. The common and intuitive explanation of the endowment effect is that owners construe giving up an object as a painful loss that counts more than money they receive in exchange, so that they demand a lot of money for the object. But if gains and losses were defined over the value of the entire transaction, owners would not be more sensitive to giving up the object than to receiving money in exchange, so no endowment effect would ensue.}

It is also crucial for our results: if gain-loss utility was defined
over total consumption utility—as would be the case, for example, in an experiment with induced values—then for any reference point the consumer’s willingness to pay for the product would be \( v \), so that the firm would set a deterministic price equal to \( v \). Second, since the gain-loss utility function \( \mu \) is the same in the two dimensions, the consumer’s sense of gain or loss is directly related to the intrinsic value of the changes in question. In Section 5 we argue that this assumption is not as crucial for our results.

Although our model does not explicitly allow for the consumer to buy other goods, it is equivalent to a specification in which the consumer spends her leftover money on a divisible alternative product, and evaluates gains and losses in the alternative product separately from gains and losses in the firm’s product. Once again, however, if the consumer integrates the gains and losses—for instance because the products satisfy very similar hedonic desires—the firm can never sell its own product more expensively than the alternative’s price, so that a different model results.

Since we assume below that expectations are rational, and in many situations such rational expectations are stochastic, we extend the utility function in Equation 1 to allow for the reference point to be a pair of probability distributions \( F = (F^v, F^p) \) over the two dimensions of consumption utility. In this case, total utility from the outcome \((k^v, k^p)\) is

\[
U(k^v|F^v) + U(k^p|F^p) = \int_{r^v} u(k^v|r^v)dF^v(r^v) + \int_{r^p} u(k^p|r^p)dF^p(r^p) .
\]

In evaluating \((k^v, k^p)\), the consumer compares it to each possibility in the reference lottery. If she had been expecting to pay either $15 or $20 for the product, for example, paying $17 for it feels like a loss of $2 relative to the possibility of paying $15, and like a gain of $3 relative to the possibility of paying $20. In addition, the weight on the loss in the overall experience is equal to the probability with which she had been expecting to pay $15.

To complete our theory of consumer behavior with the above belief-dependent preferences, we specify how beliefs are formed. We assume that they must be consistent with rationality: the consumer correctly anticipates the implications of her period-0 plans, and makes the best plan she knows she will carry through. While the formal definitions below are notationally somewhat cumbersome, the logical consequences of this requirement are intuitively relatively simple. Note that any plan of behavior formulated in period 0—which in our setting amounts simply to a strategy
of which prices to buy the product for—induces some expectations in period 0. If, given these expectations, the consumer is not willing to follow the plan, then she could not have rationally formulated the plan in the first place. Hence, a credible plan in period 0 must have the property that it is optimal given the expectations generated by the plan. Following original definitions by Kőszegi (2009) and Kőszegi and Rabin (2006), we call such a credible plan a *personal equilibrium* (PE). Given that she is constrained to choose a PE plan, a rational consumer chooses the one that maximizes her expected utility from the perspective of period 0. We call such a favorite credible plan a *preferred personal equilibrium* (PPE).

Formally, notice that whatever the consumer had been expecting, in period 1 she buys at prices up to and including some cutoff (recall that the consumer’s indifference is broken in favor of buying). Hence, any credible plan must have such a cutoff structure. Consider, then, when a plan to buy up to the price $p^*$ is credible. This plan induces an expectation $F^v(\Pi, p^*)$ of getting consumption utility $v$ from the product with probability $\Pi(p^*)$, and an expectation $F^p(\Pi, p^*)$ of spending nothing with probability $1 - \Pi(p^*)$ and spending each of the prices $p \leq p^*$ with probability $\Pr_{\Pi}(p)$. The plan is credible if, with a reference point given by these expectations, $p^*$ is indeed a cutoff price in period 1:

**Definition 2.** The cutoff price $p^*$ is a *personal equilibrium* (PE) for price distribution $\Pi$ if for the induced expectations $F^v(\Pi, p^*)$ and $F^p(\Pi, p^*)$, we have

$$U(0|F^v(\Pi, p^*)) + U(0|F^p(\Pi, p^*)) = U(v|F^v(\Pi, p^*)) + U(-p^*|F^p(\Pi, p^*)).$$  

Now utility maximization in period 0 implies that the consumer chooses the PE plan that maximizes her expected utility:

**Definition 3.** The cutoff price $p^*$ is a *preferred personal equilibrium* (PPE) for price distribution $\Pi$ if it is a PE, and for any PE cutoff price $p^{**}$,

$$E_{F^v(\Pi, p^*)}[U(k^v|F^v(\Pi, p^*))] + E_{F^p(\Pi, p^*)}[U(k^p|F^v(\Pi, p^*))]$$
$$\geq E_{F^v(\Pi, p^{**})}[U(k^v|F^v(\Pi, p^{**}))] + E_{F^p(\Pi, p^{**})}[U(k^p|F^v(\Pi, p^{**}))].$$  

(3)
The monopolist is a standard risk-neutral profit-maximizing firm, trying to maximize expected profits given the consumer’s behavior. To be able to state the monopolist’s problem simply as a maximization problem rather than as part of an equilibrium, we assume that the consumer chooses the highest-purchase-probability PPE. With this assumption, the monopolist solves

\[
\max_{\Pi} \{ \Pi(p^*) E[p|p \leq p^*] - \Pi(p^*) c \mid p^* \text{ is the highest PPE for } \Pi(\cdot) \}.
\] (4)

4 The Optimal Price Distribution

In this section, we first identify the monopolist’s optimal pricing strategy in our basic model, showing that it is consistent with the facts on supermarket pricing discussed in Section 2. We then introduce price-discovery costs, and establish that—consistent with the pricing behavior of some other retailers—if these costs are high, the monopolist chooses a sticky price.

4.1 Basic Results

Our main proposition identifies the features of the monopolist’s limit-optimal price distribution:

**Proposition 1.** Fix any \(\eta > 0\) and \(\lambda > 1\). If the firm can profitably sell to the consumer, then the profit-maximizing price distribution induces purchase with probability one. Furthermore, in that case for any limit-optimal price distribution \(\Pi(\cdot)\) there is an \(s\) satisfying \(0 < s < 1\) and a \(p(s) > p = (1 + \eta)v/(1 + \eta\lambda)\) such that (i) \(\Pi(\cdot)\) puts weight \(s\) on the interval \([p, p(s)]\), where it is continuously distributed with density \(\pi(p) = (1 + \eta\lambda)/[\eta(\lambda - 1)(v + p)]\); and (ii) \(\Pi(\cdot)\) puts weight \(1 - s\) on a single price \(p > p(s)\). The monopolist’s expected revenue is strictly greater than \(v\).

Proposition 1 says that the limit-optimal price distribution has two parts (as illustrated in Figure 1): an interval of continuously distributed low prices, and a single atomic high price. Furthermore, there is a gap between the low price interval and the price atom. Thinking of the low prices as the non-sticky sale prices and the high isolated pricing atom as the sticky regular price, this price
distribution is broadly consistent with the evidence on supermarket pricing summarized in Section 2.4.

We break down the explanation of Proposition 1 into five steps, arguing in turn that (1) the optimal deterministic price is \( v \); (2) the firm can earn more than \( v \) with a stochastic price distribution for which it is not credible for the consumer not to buy at low (sale) prices; (3) it is optimal to use variable sales prices; (4) it is suboptimal to rely solely on these “forcing” sale prices; and (5) the high (regular) price is sticky.

**Step 1.** We start by considering what the monopolist can achieve with a deterministic price \( p \). To identify the consumer’s behavior—the PPE—with such a price, we first identify the PE by solving for conditions under which the consumer is willing to follow respective plans not to buy and to buy. Suppose that the consumer had expected not to buy the product. If she buys, her consumption utility is \( v - p \), and her gain-loss utility—consisting of a gain of \( v \) in the product and a loss of \( p \) in money—is \( \eta v - \eta \lambda p \). If she does not buy, both her consumption utility and (as her outcomes conform to her expectations) her gain-loss utility are zero. Hence, she is willing to follow a plan not to buy, and therefore not buying is a PE, if and only if

\[
p > \frac{1 + \eta}{1 + \eta \lambda} \cdot v \equiv p^*.
\]

Note that in addition to saying that not buying is a PE for deterministic prices \( p > p^* \), the above considerations imply that for any price distribution, not buying for prices less than or equal to \( p^* \) is not credible.

Similar calculations show that buying at a deterministic price \( p \) is a PE if and only if

\[
p \leq \frac{1 + \eta \lambda}{1 + \eta} \cdot v \equiv p_\lambda.
\]

These observations in turn mean that there are three pertinent ranges of the monopolist’s price. For \( p > p_\lambda \), the unique PE is not to buy. For \( p \leq p_\lambda \), the unique PE is to buy. But for the range in between, the consumer’s expectations are self-fulfilling: she buys if and only if she had been

---

1. Notice that in our model, the monopolist’s cost might be higher than some of the prices it charges. This provides a non-predatory rationale for potential below-marginal-cost pricing of a single-product firm.
2. This is essentially the same analysis as that in Kőszegi and Rabin (2006, Section IV). The only difference is that unlike in Kőszegi and Rabin (2006), in the current setting there is no mixed-strategy PE because we have assumed that whenever the consumer is indifferent between buying and not buying, she buys the product with probability 1.
expecting to. Intuitively, the consumer’s expectation to get the product both generates a loss from not getting it and eliminates the loss from paying for it, so that she has a higher reservation price than if she had no expectation to get it. More generally, if the consumer expects to get the product with higher rather than lower probability, she will experience not getting it as more of a loss and paying for it as less of a loss, increasing her willingness to pay for it. This “attachment effect” (Kőszegi and Rabin 2006) generated by the expectation to buy will feature numerous times in our analysis below.

When there is a unique PE, it is also the PPE. But when there are multiple PE—for \( p \in (\underline{p}, \overline{p}] \)—the PPE is determined as the consumer’s favorite PE from the perspective of period 0. Since in each PE the consumer gets the outcome she expects, her gain-loss utility in each PE is zero. With total utility equal to consumption utility, the PPE is to maximize consumption utility—to purchase if and only if \( p \leq v \). This implies that the highest revenue the monopolist can earn with a deterministic price is \( v \).

**Step 2.** Surprisingly, due to a loss-averse decisionmaker’s dislike of uncertainty, the monopolist can get the consumer to buy at prices above \( v \) by using uncertain prices. To see why this is the case, suppose that the monopolist charges the price \( \underline{p} \) with probability \( s \) and a price \( p > \underline{p} \) with probability \( 1 - s \). Then, as we have noted above, in any PE the consumer buys at \( \underline{p} \). Given this, when would the consumer prefer a plan of buying also at \( p \)? First, consider making and following through a plan to buy only at \( \underline{p} \). If the price turns out to be \( \underline{p} \), the consumer’s consumption utility is \( v - \underline{p} \), and her gain-loss utility is \( (1 - s)\eta v - (1 - s)\eta \lambda \underline{p} \): relative to the possibility of not buying, which the consumer expected to occur with probability \( 1 - s \), buying at \( \underline{p} \) generates a gain of \( v \) in the product and a loss of \( \underline{p} \) in money. If the price turns out to be \( p \), the consumer’s consumption utility is zero, and her gain-loss utility is \( -s\eta \lambda v + s\eta p \): relative to the possibility of buying at price \( \underline{p} \), which the consumer expected to occur with probability \( s \), not buying generates a loss of \( v \) in the product and a gain of \( p \) in money. Overall, the consumer’s expected utility is

\[
s(v - \underline{p}) - \eta(\lambda - 1)s(1 - s)(v + \underline{p}).
\]  

Notice that expected gain-loss utility is negative. For an intuition, take for instance the product dimension. If the consumer gets the product, she experiences this as a gain relative to the ex-ante
expected possibility of not getting it; and if she does not get the product, she experiences this as a loss relative to the ex-ante expected possibility of getting it. Due to loss aversion, the latter feeling is stronger, so that expected gain-loss utility is negative. More generally, a decisionmaker with rational-expectations-based loss aversion dislikes uncertainty in consumption utility because she dislikes the possibility of a resulting loss more than she likes the possibility of a resulting gain (K˝ oszegi and Rabin 2007, Macera 2009, Herweg, Müller and Weinschenk 2010). In our case, the consumer faces uncertainty both in whether she gets the product and in how much she will pay, so her gain-loss disutility is proportional to $v + p$.

If the consumer makes and follows through a plan to buy at both prices, her expected utility is

\[ v - sp - (1 - s)p - \eta(\lambda - 1)s(1 - s)(p - p). \]

Once again, expected gain-loss utility is negative. But in this case, the uncertainty is only in how much the consumer will pay—not in whether she gets the product—so her gain-loss disutility is proportional to $p - p$. Comparing the above two expressions, the consumer prefers to buy at both prices if

\[ p \leq v + \frac{2\eta(\lambda - 1)sp}{1 + \eta(\lambda - 1)s}. \]

Because the strategy of buying at both prices eliminates the uncertainty in whether she will get the product, the consumer is willing to buy at a price exceeding $v$. Simple arithmetic shows that for small $s > 0$, this extra willingness to pay more than compensates the monopolist for having to charge $p$ with probability $s$.

The intuition for this last, crucial, point is that the monopolist exploits a novel type of time inconsistency that arises in our model despite a rational consumer’s attempt to maximize a single well-defined utility function.\(^6\) Notice that the expectation to buy at $p$ has a negative effect on the consumer’s expected utility: if she expects not to purchase at $p$, her utility from not purchasing at $p$ is zero; but if she expects to purchase at $p$, her utility from not purchasing at $p$—consisting of a loss of $v$ in the product and a gain of $p$ in money from comparing not purchasing to purchasing at $p$—is $-\eta\lambda v + \eta p < 0$. But when the consumer makes her purchase decision in period 1, she takes

---

\(^6\) That beliefs-based preferences can generate time-inconsistent behavior has been pointed out by Caplin and Leahy (2001) and K˝ oszegi (2009), and explored in more detail by K˝ oszegi (2009) and K˝ oszegi and Rabin (2009).
the reference point (formed in period 0) as given, and therefore ignores this negative effect. As a result, the monopolist can push the consumer’s expected utility below zero, so that it can charge an average price greater than \( v \).

**Step 3.** Having illustrated the profitability of using low “sale” prices to make not buying non-credible for the consumer, we next discuss how to choose such “forcing” sale prices. In the above example, the monopolist guarantees that the consumer buys with probability \( s \) in any PE by charging \( p \) with probability \( s \). But the monopolist can achieve the same with a higher profit by using more sale prices. In particular, consider a distribution which puts weights of \( q \) and \( s - q \) on \( p \) and \( p' > p \), respectively, where \( 0 < q < s \). Once again, not buying at price \( p \) is not credible. Moreover, we show that for a sufficiently small \( p' > p \), neither is it a PE for the consumer to buy only at price \( p \). If the consumer expected to buy only at \( p \), her consumption utility from buying at \( p' \) would be \( v - p' \), and her gain-loss utility would be \( (1 - q)\eta v - (1 - q)\eta \lambda p' - q\eta \lambda (p' - p) \): relative to the possibility of not buying, which the consumer expected to occur with probability \( 1 - q \), buying at price \( p' \) generates a gain of \( v \) in the product and a loss of \( p' \) in money; and relative to the possibility of buying at price \( p \), which the consumer expected to occur with probability \( q \), buying at \( p' \) generates no gain or loss in the product and a loss of \( p' - p \) in money. In the same situation, the consumer’s utility from not buying would be \( -q\eta \lambda v + q\eta p \): relative to the possibility of buying at price \( p \), which the consumer expected to occur with probability \( q \), not buying generates a loss of \( v \) in the product and a gain of \( p \) in money. Comparing the above two expressions, the consumer buys at price \( p' \) if

\[
p' \leq p + \frac{q\eta(\lambda - 1)(v + p)}{1 + \eta \lambda}.
\]

Since the above cutoff is greater than \( p \), if \( p' \) is sufficiently close to \( p \) it is not credible for the consumer not to buy at \( p' \). In this case, the consumer buys at both prices in any PE.

---

7 This intuition is somewhat incomplete: because paying a stochastic price generates negative expected gain-loss utility for the consumer, the fact that she has negative expected utility does not immediately imply that she has negative expected consumption utility—which is what is necessary for her to pay an average price above \( v \). For a complete intuition, therefore, we must argue that the consumer’s total disutility exceeds that from paying the stochastic price. To see this, note that if the firm charges a high price of approximately \( v \) (which is the case for small \( s \)), then the gain-loss disutility from price uncertainty is proportional to \( \eta(\lambda - 1)(v - p) \). Since \( \eta \lambda v - \eta p > \eta(\lambda - 1)(v - p) \), this is smaller than the consumer’s total disutility: facing a loss of \( v - p \) which is the consumer’s disutility from facing uncertainty of \( v - p \) in how much she will pay—is lower than facing a loss of \( v \) and a gain of \( p \)—which is the disutility the consumer imposes on herself through time-inconsistent behavior.
Intuitively, due to the attachment effect discussed above, the consumer’s realization that she will buy for a particular sale price raises her willingness to pay for the product. As a result, the monopolist needs to charge each sale price only with sufficient probability such that not buying at the next lowest possible sale price becomes non-credible. For very small $\Delta$, this distribution approximates the continuous distribution of sale prices identified in Proposition 1.

**Step 4.** While choosing sale prices to make not buying non-credible is a central part of the monopolist’s strategy, it is not optimal to make always buying the only credible plan. Suppose by contradiction that such a “forcing” distribution is optimal. By Step 2, its average price must then be greater than $v$. To get a contradiction, we argue that the consumer will still buy at all prices if the monopolist raises the highest price $p$ in the distribution to some $p' > p$ while leaving the rest of the distribution unchanged. By the definition of forcing, $p$ is such that the consumer buys at $p$ if she had been expecting to buy at prices less than $p$. Then, because the attachment effect implies that expecting to buy at $p'$ raises the consumer’s willingness to pay for the product, there is a range of $p' > p$ such that buying at all prices remains a PE (albeit not the only one). Now notice that expecting to buy at $p'$ has a positive externality on utility when buying: it eliminates losses in money and gains in the good, and since the average price is greater than $v$, the former effect dominates. This means that for $p'$ sufficiently close to $p$, the consumer prefers a plan to buy at all prices rather than only at prices below $p$, so that buying at all prices is the PPE.

**Step 5.** Since making always buying the only credible strategy is not optimal, it must be the case that the consumer prefers the plan of buying at the monopolist’s high prices rather than only at the forcing sale prices. To conclude our exposition of Proposition 1 we argue that it is optimal to choose these high “regular” prices to be sticky. Just as she dislikes uncertainty in whether she gets the product, the consumer dislikes uncertainty in the regular price: she experiences a gain if the regular price turns out to be lower than it could have been and a loss if the regular price turns out to be higher than it could have been, and due to loss aversion she feels the loss more heavily. If the regular price was variable rather than sticky, therefore, the consumer would still buy only if she was compensated with a lower average price. This creates a strong incentive for the firm to
eliminate variation in the regular price.\footnote{It is worthwhile to note why the same reasoning does not imply that the monopolist should choose a sticky sale price. Although the consumer dislikes variation in sale prices, since she buys at these prices in any PE, she cannot avoid the variation. As a result, the monopolist has no incentive to reduce variation in sale prices.}

Beyond the shape of the optimal price distribution, the observation in Step 2 that the consumer buys at an expected price exceeding \( v \) has an immediate welfare implication:

**Proposition 2.** For any \( \eta > 0, \lambda > 1, \) and \( \Delta < v - p_a \), the consumer would be better off expecting and following through a strategy of never buying than expecting and following through her actual strategy of buying at all prices.

Proposition 2 identifies a sense in which the firm’s sales are manipulative: they lead the consumer to buy the product even though she would prefer not to.\footnote{Although we model neither multi-product retailers nor the wholesaler-retailer relationship, Proposition 2 suggests that retailers may benefit less from sales than wholesalers: if welfare-reducing manipulative sales induce some consumers to avoid visiting the retailer, they lower profits from other wholesalers’ products. One would then expect wholesalers to encourage the use of sales in their contracts with downstream retailers.} Several caveats regarding this result are in order. First, the extreme version of the result—that the firm does only harm to the consumer by selling to her—clearly relies on our assumption that the firm knows the consumer’s preferences perfectly. Consumers with much higher valuation than the range of possible prices would clearly be better off buying than not buying. Nevertheless, Proposition 4 below shows that even with consumer heterogeneity, some consumers who buy with positive probability would be better off making and following through a plan of never buying. Second, it matters what the consumer would do with the money if she did not buy from this firm. Given that we assume linear consumption utility in money, the implicit assumption in our model is that the consumer would spend her money on an alternative divisible product which is available on the market at a deterministic price. But if she would be manipulated into buying something else from another firm, she might be better off buying from this firm. Third, alternatives to our rational-expectations assumption, such as that discussed in Section 5.3, might affect the welfare implications of the consumer’s behavior—even if they do not qualitatively change the optimal price distribution.
4.2 Costly Price Discovery

In this section we analyze pricing in our model when the consumer does not automatically observe whether the product is on sale. Comparing this version to the basic one above will allow us to identify a potentially important difference between types of retailers implied by our theory.

Suppose that to see the price realization in period 1, the consumer must pay a monetary or effort cost $\phi$ satisfying $0 < \phi < v$. We assume that gain-loss utility in the price-discovery cost is evaluated separately from the product and money. As will be apparent from our argument, the results would be the same if the cost was on the same dimension as money. Note that since the consumer decides whether to pay $\phi$ before she knows the price realization, this decision cannot be made contingent on whether she ends up purchasing. We assume that the monopolist must choose a price distribution over the non-negative reals.

Proposition 3 identifies the features of the optimal price distribution in this case:

**Proposition 3.** For any $\eta > 0$, $\lambda > 1$, there is a $\phi^*$ satisfying $0 < \phi^* < p$ such that:

I. If $\phi < \phi^*$ and the firm can profitably sell to the consumer, then for any limit-optimal price distribution $\Pi(\cdot)$ there is an $s$ satisfying $0 < s < 1$ and prices $p' = (1 + \eta + \eta(\lambda - 1)\phi/p)v/(1 + \eta\lambda), p''$ satisfying $p < p' < p'' < p_{reg}$ such that (i) $\Pi(\cdot)$ puts weight $\phi/p$ on zero; (ii) $\Pi(\cdot)$ puts weight $s$ on the interval $[p', p'']$, where it is continuously distributed with density $\pi(p) = (1 + \eta\lambda)/(\eta(\lambda - 1)(v + p))$; and (iii) $\Pi(\cdot)$ puts weight $1 - s - \phi/p$ on $p_{reg}$. The monopolist’s expected revenue is strictly greater than $v - \phi$.

II. If $\phi > \phi^*$ and the firm can profitably sell to the consumer, then the limit-optimal price distribution puts probability one on $v - \phi$.

The first part of Proposition 3 says that if $\phi$ is relatively small, the limit-optimal price distribution is very similar to that with no price-discovery costs (Proposition 1), with one crucial difference: the monopolist charges a price of zero with some probability. Intuitively, the “free sample” is part of the monopolist’s scheme to make not buying non-credible for the consumer. If the consumer had expected not to buy, she will pay $\phi$ to check the price realization only if she has a chance of getting the product at a price below $p$, and in this range she saves at least $\phi$ in expectation relative to $p$. 

19
The profit-maximizing way to give away these savings is to sometimes offer free samples. More interestingly, if \( \phi \) is relatively large, the limit-optimal price distribution is deterministic—the price is completely sticky. This is easiest to see when \( \phi > p \). In that case, if the consumer had expected not to buy, the monopolist cannot get her to check the price realization even if it charges a price of zero with probability one. Hence, manipulating the consumer into buying against her will is impossible, so that there is no point in using sales. But going further, even if the firm can manipulate the consumer into buying against her will, for a range of \( \phi \) it is suboptimal to do so. Intuitively, if the consumer expects to pay lower prices for the product, she experiences paying higher prices as more of a loss, and is therefore less willing to buy at those prices. Hence, if the firm sometimes offers very low sale prices, it must lower its prices at the high end as well. This makes using stochastic prices relatively less attractive.

While we are not aware of a systematic empirical analysis on the relationship between price-discovery costs and the frequency of sales, our theoretical prediction on this relationship provides a potential explanation for why sales are much more common at supermarkets than at many other kinds of retailers. A typical consumer at a supermarket visits to make some planned purchases of supplies, but is also willing to consider additional products to buy. Once at the supermarket, the consumer’s cost of learning the price realization of many additional products is very close to zero, so that the optimal price distribution involves stochastic sales. In contrast, for many other types of goods—e.g. movies—a typical consumer will not visit the retailer unless she buys some core product sold by the retailer. Without being at the retailer by default, the consumer’s price-discovery costs

10 To see this, suppose that the expected price below \( p \) is positive. The firm could then redistribute the same weight on zero and \( p \) leaving the average price at or below \( p \) unaffected, maintaining the property that in any PE the consumer pays the price-discovery cost and buys at both of these prices. Furthermore, since \( p \) is the highest price at which the consumer prefers to buy in period 1 if she had expected to buy with probability zero, but she actually expects to “purchase” with positive probability at a price of zero, the attachment effect implies that in any PE she strictly prefers to buy at \( p \) in period 1. Hence, the firm can move the atom at \( p \) to a slightly higher price, while still maintaining the property that in any PE the consumer pays the price-discovery cost and buys at both prices. Since this does not undermine the firm’s ability to sell at higher prices, it increases profits.

11 Although our model provides a novel potential explanation for why firms sometimes offer free samples of their products, this prediction is not particularly robust to natural variations of our model. Since there is only one consumer who buys at most one item, our model presupposes that the firm’s free offer cannot be exploited using resale or storage, and will not be used by consumers who do not intend to buy at higher prices. If these additional considerations are important, the firm will put the atom at a higher price, or switch to a deterministic pricing strategy as explained below.
are non-trivial, so that the optimal price distribution is deterministic.

5 Extensions and Modifications

To further illustrate the forces driving our results, as well as to explore additional issues, in this section we investigate various extensions and modifications of our model.

5.1 Heterogeneous Consumers and Cost Uncertainty

We have emphasized throughout the paper that our model predicts a sticky (atomic) regular price. But this result may seem to rely on our assumption that the firm faces a deterministic environment, an assumption that would lead a classical monopolist to choose a sticky price also. Hence, we analyze a variant of our model in which consumers have heterogeneous consumption valuations $v$ and the firm faces cost shocks—assumptions that in a classical reference-independent setting would generate price variation. While we cannot solve for the fully general optimal price distribution, we illustrate the robustness of the stickiness of the regular price (as well as the optimality of stochastic prices and the variability of sale prices) by restricting the space of price distributions from which the firm can choose.

Specifically, suppose that there is a continuum of consumers whose consumption utility $v$ is uniformly distributed on the interval $[0, v_H]$. The firm’s marginal cost is continuously distributed on the interval $[c_L, c_H]$, with $0 < c_L < c_H < v_H$. We think of the firm as choosing the price distribution the consumer faces; given any chosen price distribution, it is then optimal to charge lower prices in lower-cost states.

---

12 The internet has, of course, made it much less costly to find out information about retailers, likely decreasing price-discovery costs. For example, if a movie theater were to use the sales-and-regular-prices strategy, it could post up-to-date price information on the internet. Even so, a consumer’s price-discovery cost would clearly be higher than at a supermarket, and could still be non-trivial relative to a marginal consumer’s value of the product. Furthermore, because a marginal consumer is made worse off when the firm uses a sales-and-regular-prices strategy, she would not want to make it easy for herself to check the price realization.

13 The condition that $c_H < v_H$ ensures that the firm can profitably sell to the consumer. For example, if the firm charges the price $(v_H + c_H)/2$ with probability one, then (applying our analysis from Section 4.1 Step 1) all consumers with valuation greater than this price buy the good, so that for any realized marginal cost the firm earns positive profits.
We first argue that the monopolist can still make higher profits with a stochastic price distribution than with a deterministic one. As we have shown (Section 4.1, Step 1), for a deterministic price \( p \) consumers maximize consumption utility, buying the product if and only if \( p \leq v \). Hence, for deterministic prices the monopolist faces a classical downward-sloping demand curve. Suppose a profit-maximizing deterministic price is \( p \). Consider instead the optimal price distribution from Section 4.1 for the consumer who has \( v = p \). It is easy to see that consumers with consumption utility \( v < p \) will not buy the product for any price, and consumers with consumption utility \( v \geq p \) will buy the product at all prices. Hence, the firm sells the same amount for a higher average price and the same average cost, increasing profits.

To illustrate the robustness of our other basic findings, we restrict attention to price distributions in which the prices \( p_L - \alpha_L, p_L + \alpha_L, p_H - \alpha_H, \) and \( p_H + \alpha_H \) are charged with probabilities \( s/2, s/2, (1 - s)/2, \) and \( (1 - s)/2, \) respectively. Constrained by the exogenous bound \( \bar{\alpha} > 0 \), the firm chooses \( p_L, p_H > p_L + 2\bar{\alpha}, \alpha_L \) and \( \alpha_H \) satisfying \( 0 \leq \alpha_L, \alpha_H \leq \bar{\alpha} \), and \( s \in [0, 1) \). While restrictive, this class of price distributions allows us to reconsider each of the features of the optimal price distribution we have found above. Whether the monopolist chooses \( s > 0 \) answers whether it would like a distribution of the sales-and-regular-prices structure; whether it chooses \( \alpha_L > 0 \) answers whether it would like variable sale prices; and whether it chooses \( \alpha_H > 0 \) answers whether it would like a variable regular price.

As a point of comparison, consider first what happens when consumers are not loss averse. Fixing any \( \bar{\alpha} > 0 \), if \( c_H - c_L \) is positive but sufficiently small, the firm sets \( s = 0 \) and \( \alpha_H > 0 \)—it does not engage in a strategy of significantly different sales and regular prices, but it does respond to small cost shocks by varying its price a little bit. And if \( c_H - c_L = 0 \), then the firm sets \( s = 0 \) and \( \alpha_H = 0 \)—it chooses an atomic price. In contrast, with loss-averse consumers the firm engages in a sales-and-regular-prices strategy and uses variable sale prices whether or not \( c_H - c_L > 0 \), yet it chooses not to respond to small cost shocks by varying its regular price:

**Proposition 4.** Fix any \( \eta > 0, \lambda > 1, \) and \( c_L > 0 \). Then, if \( c_H - c_L \) and \( \bar{\alpha} \) are sufficiently small, the optimal price distribution has \( s > 1, \alpha_L = \bar{\alpha}, \) and \( \alpha_H = 0 \). A positive measure of consumers would be strictly better off making and following through a plan of never buying than making and
following through their PPE plan.

The reason that the firm chooses a regular-prices-and-sales strategy is the same as above and as in our basic model: to manipulate some consumers into buying at an average price above their valuation. This pricing strategy of course makes the affected marginal consumers worse off relative to following a strategy of never buying. The reason that the optimal sale prices are variable is also the same as in our basic model: once a consumer knows she will buy at a price \( p \), it is not credible for her to forego buying at slightly higher prices.

The most interesting novel prediction in Proposition 4 is that even with some cost uncertainty—and despite the fact that the overall price distribution is highly uncertain—the monopolist does not respond to small cost shocks by varying the regular price. Intuitively, loss-averse consumers respond very differently to variation in the regular price than do classical consumers. In a classical model, the monopolist sells more at lower prices, so varying the price allows it to save on production costs by shifting more sales into lower-cost states. With loss aversion, however, small amounts of regular-price variation decrease demand at both the high and the low regular price. Because it eliminates uncertainty in whether she will get the product at the cost of paying only a slightly higher price, for small \( \alpha_H \) any consumer who buys at price \( p_H - \alpha_H \) also buys at price \( p_H + \alpha_H \). But because consumers dislike uncertainty in the regular price, they are less likely to choose to buy at these prices if \( \alpha_H > 0 \) than if \( \alpha_H = 0 \). As a result, a slightly variable regular price reduces demand without shifting more production into lower-cost states, so that for sufficiently small cost shocks the firm does not vary the regular price.

To illustrate the robustness of Proposition 3 we also briefly discuss what happens in the above setting when consumers have high price-discovery costs. Then, so long as \( c_H \) and \( c_L \) are not too different, the optimal price distribution is deterministic: \( s = 1 \) and \( \alpha_H = 0 \). The intuition is similar to that of Proposition 3 because consumers cannot be manipulated into buying they dislike, the firm must induce consumers to want to buy. And since loss-averse consumers strongly dislike price variation, the firm benefits from eliminating price variation altogether.
5.2 Competition

Our main analysis focuses on the case of a monopolistic retailer. While the general question of how competition affects pricing is beyond the scope of the current paper, we discuss three simple forms of competition. First, we consider perfect ex-ante competition for consumers, as for example when consumers decide which supermarket or restaurant to frequent. Suppose that there is a mass of consumers whose consumption value is distributed continuously on the interval \([0, v_H]\), with positive density everywhere. Two retailers simultaneously commit to a price distribution. After observing the distributions, consumers decide which retailer to visit, and form expectations about their consumption outcomes. We assume that if indifferent consumers choose one of the two retailers randomly with equal probability. Finally, a price is drawn from each retailer’s price distribution, and consumers decide whether to buy at their previously chosen retailer’s price. The two retailers have identical cost distributions uniformly distributed on the interval \([c_L, c_H]\) with density \(d\), where \(0 < c_L < c_H < v_H\). Then:

\[\text{Proposition 5. Fix any } \eta > 0, \lambda > 1, \text{ and } (c_L + c_H)/2 > 0. \text{ If } d \text{ is sufficiently large, then for any } \Delta > 0 \text{ the unique equilibrium with ex-ante competition is for each firm to choose the deterministic price } (c_L + c_H)/2.\]

Proposition 5 says that if the firms’ costs are sufficiently densely distributed, the unique equilibrium is for each firm to choose the deterministic price equal to average cost. Intuitively, because a loss-averse consumer dislikes price uncertainty, to attract her a firm has an incentive to eliminate price variation. This strategy is reminiscent of some retailers’ (most notably Walmart’s) promise to have “Everyday Low Prices” rather than fluctuating prices. More systematically, Proposition 5 may explain why restaurants—often facing stiff local competition for patrons—do not seem to have frequent sales like supermarkets (MacDonald and Aaronson 2006). The proposition also predicts that supermarkets operating in a very competitive market for new consumers will have a policy of charging stable prices and not having random sales.

Note that sticky pricing is not an equilibrium in the above model when consumers have classical

\[\text{\footnote{The assumption that costs are uniformly distributed simplifies our calculations, but is not crucial: sticky pricing would result so long as the cost distribution is sufficiently dense everywhere on } [c_L, c_H].}\]

24
reference-independent preferences, even if these consumers are risk-averse with respect to the surplus from the transaction or the price to be paid for the product. If a firm charges the deterministic price equal to average cost, its competitor can profitably deviate by offering lower prices when its costs are lower, attracting some consumers whose value is below the average cost.

As a second form of competition, we consider ex-post competition for consumers, where consumers decide which retailer to buy from after seeing the price realizations. In that case, as with reference-independent utility, perfect competition yields marginal-cost pricing. Hence, if costs are volatile, so too will be prices. This situation is consistent, for instance, with the market for economy-class airline tickets, in which many consumers shop around among a number of very competitive retailers before each purchase—and in which prices are indeed very volatile.

Summarizing, our theory therefore predicts vastly different effects of competition on price volatility depending on the exact form of competition. Whereas a monopolist holds random sales even if it faces a deterministic environment, firms engaged in ex-ante competition keep their prices sticky even if costs vary somewhat. In contrast to both of these pricing strategies, firms engaged in ex-post competition vary prices together with costs.

Finally, we discuss a form of imperfect competition. Suppose the monopolist faces a competitive fringe: there is a competitive industry producing a substitute product that has a lower consumption value \( v_f < v \) on the same dimension as the monopolist’s product, the consumer is interested in buying at most one of the products, and she decides which one to buy after seeing both prices. The competitive fringe charges a low price \( p_f \leq (1 + \eta) v_f / (1 + \eta \lambda) \). In this case, whatever the consumer had expected, she prefers to buy the fringe’s good to not consuming. Hence, in any PE she buys one of the products, getting intrinsic utility of at least \( v_f \). Therefore, if the higher-value product is priced at \( p_f \), the consumer buys that product, so that the firm never chooses a price below \( p_f \).

This implies that in any PE, the consumer gets consumption utility from the product of at least \( v_f \) and pays a price of at least \( p_f \). As a result, the firm’s problem can be thought of as choosing the distribution of the price premium \( p - p_f \) it charges for the incremental consumption value \( v - v_f \). Therefore, the optimal price distribution is the same as that of a monopolist who sells a product of value \( v - v_f \), shifted to the right by \( p_f \)—it has the same shape and probability of sales as the
optimal price distribution in our basic model, but it is more compressed.

5.3 Further Extensions and Modifications

In our basic model, we have taken the representative consumer’s consumption value \( v \) to be deterministic. Suppose instead that \( v \) is uncertain. We can distinguish two cases, depending on whether the consumer knows \( v \) in advance (in period 0). If she does, then (although we have not analyzed such a model in detail) the same forces as with cost uncertainty are likely to operate, so that a qualitatively similar price distribution likely results. If the consumer does know \( v \) in advance, then for each \( v \) the monopolist chooses the optimal price distribution we have derived for that \( v \). For example, as we have discussed, if price-discovery costs are high our theory predicts a sticky price for each \( v \). But if \( v \) changes in a way that is predictable by consumers in advance, the optimal sticky price changes. This prediction is consistent with matinees in movie theaters and cyclical sales of many products for which the sale price is also sticky. From the perspective of our model, therefore, these cyclical sales are more appropriately viewed as changes of the entire price distribution than as shifts within a single price distribution. At the same time, our model does not explain why prices do not seem to change in response to some other predictable changes in the demand.

Our model assumes that when forming plans in period 0, the consumer can choose any plan at no immediate cost, in a sense starting from a reference-free position. It is plausible, however, that the consumer enters the marketplace in period 0 with some initial (and not necessarily rational) expectations already in her mind, and changing these expectations generates gain-loss utility in period 0. The initial expectations can be thought of as what she expects before she makes rational plans regarding the current purchase. In this alternative theory, the definition of PE is the same as in our basic model, and the PPE is the PE that maximizes total expected utility in periods 0 and 1 (see Kőszegi and Rabin 2009, Web Appendix). We have shown that for any initial expectations by the consumer, the limit-optimal price distribution is still some combination of continuously distributed sale prices and a sticky regular price. One of these two components of the price distribution, however, could be empty. Most interestingly, if the consumer enters the market expecting to get the product and pay money, and the weight on her period-0 gain-loss utility is relatively high, then
the optimal price distribution is deterministic. Intuitively, since the consumer expects to buy the product to begin with and finds this expectation painful to give up, it is not necessary to manipulate her into buying. Hence, the sales region in the price distribution loses its purpose. Assuming that a large share of consumers initially expects to buy products by popular brands, this result predicts that such brands should not have sales. Indeed, based on casual observation it seems that some brands (e.g. Apple) rarely hold sales.

Following Kőszegi and Rabin (2006), our model assumes that the loss-aversion parameters \( \eta \) and \( \lambda \) are the same in the product and money dimensions. As argued by Novemsky and Kahneman (2005) and Kőszegi and Rabin (2009), however, it might be that reference dependence and loss aversion are weaker in the money than in the product dimension. To capture this, suppose that the weight on gain-loss utility is \( \eta^v \) in the product dimension and \( \eta^p \leq \eta^v \) in the money dimension, with the coefficient of loss aversion \( \lambda \) still being the same in the two dimensions. We argue that unless \( \eta^p \) is much lower than \( \eta^v \), this modification does not fundamentally change our results.

Assuming \( \eta^p < \eta^v \) in fact strengthens the logic behind the optimality of stochastic prices. For an illustration, consider the extreme case of \( \eta^p = 0 \). Then, not buying is a credible plan if and only if prices are strictly greater than \( \bar{p} \equiv (1 + \eta^v)v \), so \( \bar{p} \) is both the highest price at which the consumer buys if she had not been expecting to do so, and the highest deterministic price at which she buys. In addition, as in our basic model, if the consumer had been expecting to buy at price \( \bar{p} \) with some probability, due to the attachment effect she buys at slightly higher prices as well. Hence, a stochastic price dominates a deterministic one.

While the monopolist wants to use a stochastic price distribution and variable sale prices for any \( \eta^p \leq \eta^v \), it does not want to have a regular price if \( \eta^p \) is much lower than \( \eta^v \). Intuitively, as we explain in Step 4 of our discussion of Proposition 1, the consumer’s motive to buy at a high regular price \( p' \) is that this turns buying into more of a gain in money (relative to paying \( p' \)) rather than a loss in money (relative to paying nothing). But if the consumer's reference-dependent utility in

\[15\] In Kőszegi and Rabin (2009), in particular, the decisionmaker derives gain-loss utility from changes in beliefs about present and future consumption utility, with news about future consumption utility less heavily felt. Then, since money paid for an item typically impacts only future consumption, the gain-loss utility from monetary outlays is lower than for products to be consumed soon. In this case, a consumer’s behavior might be captured in a reduced-form static model by assuming that she has a lower \( \eta \) in the money than in the product dimension.
money carries little weight, she does not care much about this effect, so that she is not willing to buy at a high regular price. In this case, the optimal price distribution is a distribution akin to the sale prices in our basic model above. If $\eta_p$ is not much lower than $\eta^v$, however, the same logic as in our basic model holds, and the optimal price distribution includes an isolated atom at the top.\footnote{Straightforward (but long) arithmetic shows that a sufficient condition for this is $\eta_p \lambda + (\eta_p)^2(\lambda^2 - 1) \geq \eta^v$.}

An interesting possibility arises in our model when the firm’s opportunity cost of delivering the product is sometimes greater than the highest possible price.\footnote{This could occur either because the firm itself faces high costs, or because it has another consumer with high valuation.} In a classical setting, the firm would not sell to the consumer in these contingencies. But in our theory, not getting the product in some states reduces the consumer’s willingness to pay in other states, so that the monopolist may commit to selling even in situations in which it makes losses from doing so.

For both analytical convenience and methodological discipline, we have followed K˝ oszegi and Rabin (2006) as well as classical economic methodology in assuming that the consumer fully understands the firm’s pricing strategy and immediately solves for the PPE. While we are not aware of and have not worked out a full formal model, we conclude this section with an intuitive discussion of how a less-than-fully-rational consumer might respond to the firm’s price distribution. For repeated decisions, in fact, it seems that far less than full rationality is sufficient for the consumer to eventually play PPE, and hence to give the firm similar incentives as when it faces a fully rational consumer.\footnote{The type of consumer behavior in our discussion below is consistent with the phenomenon of projection bias (Loewenstein and Schkade 1999, Loewenstein, O’Donoghue and Rabin 2003) as applied to expectations-based reference-dependent preferences: the consumer does not fully take into account that if she changes her plans or expectations, her future preferences will change also.}

Suppose, for instance, that the consumer initially comes into the market not expecting to buy the product or spend money, and she repeatedly faces purchase decisions with prices drawn from a distribution like the monopolist’s optimal price distribution in the fully rational model above. Whenever the price is $p$, she will end up purchasing the product. She will then learn that she sometimes gets the product, and build this into her expectations. If she sees the price $p$ sufficiently often, therefore, the attachment effect will mean that she buys at a slightly higher price as well. Continuing this logic further, she will buy at all sale prices. Then, if she understands at least partly that a plan to buy at the regular price will reduce the sense of loss from paying money,
she will also be willing to buy at a high regular price.\footnote{While some forms of less-than-full rationality might not affect the qualitative predictions regarding the optimal price distribution, they might still affect the welfare implications of our model. In particular, the above type of consumer does not realize that once she incorporates spending money into her expectations, such spending will no longer feel like a loss. Because she therefore overattends to the possibility of losing money, manipulating her into buying using a stochastic price distribution might be (in contrast to Proposition\ref{th:main}) welfare-improving.}

6 Related Literature

In this section, we discuss the literatures most closely related to our paper beyond the key evidence on pricing summarized in Section\ref{sec:pricing}.

6.1 Loss Aversion

Loss aversion is a natural explanation for the endowment effect, small-scale risk aversion, and other widely observed patterns in individual behavior, and seems to contribute to consumer behavior in the marketplace.\footnote{\cite{Kahneman1990, Kahneman1991}, for instance, find in a series of experiments that randomly assigned owners of an object value it more highly than non-owners, presumably because owners construe giving up the object as a loss. In addition, as argued by Rabin (2000), Rabin and Thaler (2001), Barberis, Huang and Thaler (2006), and other researchers, the most significant source of aversion to risk over modest stakes is loss aversion. And research in marketing suggests that consumers are loss averse in their evaluation of market prices (Erickson and Johansson 1985, Winer 1986, Kalwani and Yim 1992), with their “reference price” determined at least partly by expectations (Jacobson and Obermiller 1990).} Beyond this extensive evidence on the general phenomenon of loss aversion, more recent evidence also lends support to Kőszeig and Rabin’s (2006) particular, expectations-based, model of reference-dependent preferences and loss aversion. In Abeler, Falk, Götte, and Huffman’s (forthcoming) experiment, subjects work on a boring task for a piece-rate, and can choose when to stop. When they are done, a coin flip determines whether they receive what they earned or a predetermined amount, set to be 3.50 Euros for one group and 7.00 Euros for another group. A significant number of subjects stop working when they earned exactly the predetermined amount, suggesting that this expected amount became (part of) their reference point for earnings. In a simple exchange experiment, Ericson and Fuster (2009) find that subjects are more likely to keep an item they had received if they have been expecting a lower probability of being able to exchange it, consistent with the idea that their expectations affected their reference point.\footnote{In an alternative experiment, Ericson and Fuster (2009) find that subjects are willing to pay 20-30 percent more for an object if they had expected to be able to get it with 80-90% rather than 10-20% probability, consistent with...}

\footnote{20}

21

29
Crawford and Meng (2008) propose a model of cabdrivers’ daily labor-supply decisions in which cabdrivers have rational-expectations-based reference points (“targets”) in both hours and income. Crawford and Meng show that by making predictions about which target is reached first given the prevailing wage each day, their model can reconcile the controversy between Camerer, Babcock, Loewenstein and Thaler (1997) and Farber (2005, 2008) in whether cabdrivers have reference-dependent preferences.

There are also several papers investigating firm pricing with consumer loss aversion. In Heidhues and Kőszeigi (2008), we consider a model of oligopolistic competition with differentiated products, and show that due to consumer loss aversion, demand is more elastic at higher than at lower market prices, leading firms to reduce or eliminate price variation. This can explain why competitors often sell differentiated goods at identical prices—even in environments that are not perfectly symmetric. In earlier unpublished work (Heidhues and Kőszeigi 2005), we have argued that consumer loss aversion generates price stickiness, but could not simultaneously explain the stickiness of regular prices and the prevalence of variable sales. Using a somewhat different, sampling-based, model of how consumers form reference points in money and ignoring loss aversion in the product dimension, Spiegler (2010) replicates Heidhues and Kőszeigi’s (2005) main finding of price stickiness.

---

22 The plausibility of Kőszeigi and Rabin’s (2006) framework as a generally applicable model of consumer behavior is also bolstered by theoretical applications that explain puzzles in a number of important economic settings outside pricing, especially contracting. Because loss-averse decisionmakers strongly dislike variation in monetary outcomes, the theory often predicts less sensitivity of contracts to information than classical models. Using this basic insight, Herweg et al. (2010) show that the optimal way to provide incentives to exert effort while minimizing wage variation is often to use a “bonus contract” consisting of two possible wage levels. Based on a dynamic extension of Kőszeigi and Rabin (2006) proposed by Kőszeigi and Rabin (2009), Macera (2009) demonstrates that principals will often use future payments to generate current incentives, thereby eliminating variation in the current wage. And Herweg (2010) finds that flat-rate contracts are often the optimal two-part tariffs for services even when firms have positive marginal costs.

23 Karle and Peitz (2009) qualify Heidhues and Kőszeigi’s (2008) prediction of reduced price variability by showing that in some asymmetric duopolistic environments—specifically, when consumers observe prices but not how much they will like each product before their expectations-based reference point is set—consumer loss aversion can actually increase price differences.

24 Deviating from the expectations-based model, Zhou (2009) assumes that consumers take the first or most prominent price they see as the reference point for money outlays. Because the leading firm benefits a lot from having a lower price than its competitor and is hurt less by having a higher price, it has an incentive to avoid charging the same price, so that in a simultaneous-move game it sets a random price.
6.2 Classical Theories of Pricing

In this section, we discuss other theories that explain some of the same price patterns as our model. While we identify other differences below, the most important difference is that (at least without additional assumptions) none of these theories provide a robust micro-founded explanation for the combination of facts we have emphasized: that (i) at supermarkets, the regular price is sticky, while there are frequent sales with variable prices; and (ii) at many other types of retailers, sales are less common and prices are simply sticky. In particular, previous theories have explained all or part of the patterns in (i), but not the combination together with (ii). Furthermore, they do not make many of the additional predictions in Section 5 on the effects of competition and other forces.

We begin with papers that can account for all the patterns in point (i) above even in an uncertain environment that generates incentives to change regular prices. Kehoe and Midrigan (2008) assume that there are two distinct kinds of prices, regular prices and sale prices, and that there is both a menu cost associated with changing the regular price, and a different and lower menu cost associated with having an item on sale. Then, the regular price is sticky because it is costly to change, but sale prices are not sticky because (conditional on having a sale) they are costless to change. While Kehoe and Midrigan’s theory makes many of the same predictions as ours and can be incorporated tractably into macroeconomic theories for simulating the effects of monetary policy and other questions, it leaves unanswered why there would be two different kinds of prices with different menu costs. Our theory can be thought of as providing micro-foundations for these reduced-form assumptions. These micro-foundations allow us to explain without additional assumptions why many other retailers have much less frequent sales, and to make further predictions on the effect of competition and other forces.

Nakamura and Steinsson (2009) analyze a repeated price-setting game between a monopolist with privately known cost and a consumer with habit formation. Because the consumer is more willing to consume the firm’s product and develop a habit if she believes future prices will be low, the monopolist would like to commit to relatively low future prices. As a result, the monopolist’s

---

25 Some models of sales we discuss below make this set of predictions in a deterministic environment. To our knowledge, no previous theory can account for all the patterns in point (i) in both deterministic and uncertain environments, as ours does.
favorite Markov-perfect equilibrium is one in which it never selects prices above a price cap. At
the cap the price is unresponsive to cost, but below the cap the price is fully responsive to cost.
While Nakamura and Steinsson (2009) do not analyze this possibility, it seems that there could well
be higher-profit non-Markov equilibria in which the firm compensates consumers for high current
prices by charging lower prices in the future. In addition, because the price distribution is essentially
the distribution of short-run profit-maximizing prices censored at the price cap, their model does
not naturally predict a gap between the regular price and sale prices, as our model does.

Many other papers can account for part of the patterns in point (i) above, but without additional
assumptions cannot explain why some retailers have sales and others do not. There is a considerable
industrial-organization literature investigating why firms engage in sales. The most important and
most common explanation is based on firms’ incentive to price discriminate between groups of
consumers. In Conlisk, Gerstner and Sobel’s (1984) model of a durable-goods monopolist, for
example, a new cohort of heterogeneous consumers enters the market in each period, with each
consumer deciding whether to buy the good immediately or after some delay. In most periods
the monopolist sells to high-valuation buyers only, but in some periods it lowers its price to sell
to the accumulated low-valuation consumers. Intertemporal-price-discrimination models clearly
capture a realistic and important feature missing from our model, and in this sense we view them
as complementary to our theory. But these models do not explain the stickiness of the regular price
for most products or the existence of random sales for perishable goods with stable demand and
cost characteristics.

26 Conlisk, Gerstner and Sobel (1984) assume that high-valuation consumers purchase immediately unless the
price exceeds their value. Sobel (1984, 1991) relaxes this assumption and shows that stationary equilibria still involve
price cycles, while a folk-theorem result obtains for nonstationary equilibria. Pesendorfer (2002) shows that an
intertemporal-price-discrimination model with storage by consumers matches pricing and consumer behavior in the
market for ketchup quite well.

27 One feature of intertemporal-price-discrimination models that distinguish them from our theory is that they
predict higher profits at low than at high prices. Intuitively, a low price today decreases future profits by inducing
some consumers to buy now rather than later, so that a firm is willing to set such a low price only if compensated
by immediate profits. We are not aware of much evidence on profits in sale periods relative to regular-price periods.
One suggestive paper is Slade (1999), who investigates the prices set for saltine crackers by grocery stores in a small
US town. In her data, prices change infrequently and often by large amounts. She allows own past prices to have
either a negative or a positive effect on current demand. The negative effect allows for the price discrimination effects
that we have discussed in the sales literature. The positive effect is meant to capture a stock of goodwill, which could
arise through “consumer habit formation, product awareness, or brand loyalty.” In her empirical implementation,
she finds evidence that goodwill, i.e. low past prices, increase current sales.
There is also a set of models in industrial organization in which the oligopolistic environment leads firms to play mixed strategies.\footnote{See, for example, Shilony (1977), Varian (1980) Gal-Or (1982), Davidson and Deneckere (1986), Baye, Kovenock and de Vries (1992).} In all of these papers, each firm is left with a “captive” group of consumers who will not buy from a cheaper rival, and a “non-captive” group for which firms engage in price competition. In equilibrium, firms randomize between charging the monopoly price for the captive consumers and competing for the non-captive consumers. These theories either do not predict a sticky regular price, or this prediction is not robust to cost shocks when the demand of captive consumers is downward-sloping.

7 Conclusion

While our model provides a potential explanation for a number of pricing patterns, there are some patterns it cannot convincingly explain. For instance, at many establishments Persian rugs and furniture seem to be perpetually “on sale” from an essentially fictitious “regular price” that is almost never charged. For these products, consumers are unlikely to know the price distribution, and the perpetual-sale strategy probably aims to manipulate consumers’ perceptions about typical prices and quality. Similarly, given that volume tends to be much higher during sales than when the regular price is charged, it is likely that the storage motive on the part of consumers and intertemporal price discrimination on the part of firms plays an important role in sales. An important agenda for future research is to investigate how loss aversion interacts with these other forces. For instance, it seems that loss-averse consumers’ dislike of running out of the product or paying a lot for it could strengthen the storage motive.

References


### A Reputation-Based Microfoundations for Commitment

In this section, we provide foundations for our commitment assumption based on the methods of Fudenberg and Levine (1989). Suppose that the monopolist lives for infinitely many periods, and has a discount fact of $\delta < 1$. In each period $t$, the monopolist interacts with a new consumer in the following way. At the end of period $t - 1$, the consumer forms expectations regarding her consumption outcomes. In period $t$, she observes the monopolist’s price and makes a purchase decision. When forming expectations at the end of period $t - 1$, the consumer observes the monopolist’s pricing distribution in all periods 1 through $t - 1$. While extreme, the assumption that the consumer can observe the monopolist’s strategy rather than just the realized price captures the idea that consumers know a lot about a typical large retailer’s pricing strategy: they can observe the price for many products and in principle at many stores. We have not investigated whether and how our results can be extended, using the methods of Fudenberg and Levine (1992), to a situation where consumers only observe realizations of the monopolist’s mixed strategy.
For the purposes of our microfoundation, we make the following modifications to our setup of Section 3, which do not affect the logic of our main results. First, the key assumption to the reputation model is that there is a positive probability that the monopolist is a “crazy” type who does not maximize discounted expected profits, but plays according to a fixed strategy. For any $\epsilon > 0$, let $\Pi^*_\epsilon$ be an optimal committed strategy shifted to the left by $\epsilon$. We assume that with probability $\gamma > 0$ the monopolist plays $\Pi^*_\epsilon$ in every period, and with probability $1 - \gamma$ it is a rational type who maximizes its expected discounted profits.\footnote{The reason we specify the crazy type to play a strategy close to the optimal committed strategy rather than the optimal strategy itself is that the optimal strategy has a “knife-edge” feature: every sale price is charged with just enough probability for the consumer to buy at the next highest price. With such knife-edge strategies, in order for the consumer to make a plan to buy, it is not sufficient for her to be almost convinced that the monopolist is the crazy type.}

Second, we assume that in order to observe the monopolist’s price realization in period $t$, the period-$t$ consumer must pay a small cost $\phi > 0$. This assumption ensures that if the monopolist reveals itself to be a rational type, there is an equilibrium in which all consumers stay away from the store for the rest of the periods. This assumption modifies the optimal committed price distribution a little bit, as the lowest sale price must compensate the consumer for paying $\phi$. But other than this, the properties of the optimal committed price distribution remain the same.

Let $\pi^*$ be the expected per-period profits with the optimal committed distribution. Proposition 6 below says that if $\delta$ is sufficiently close to 1 and the crazy type’s strategy is close to the optimal committed strategy, the monopolist will play a strategy close to an optimal committed strategy in most of the periods. To state the proposition, we introduce the following notation. For any $\alpha > 0$, let $P_\alpha$ be the set of pricing strategies by the firm that earn profits within $\alpha$ of $\pi^*$. We note that as $\alpha \to 0$, the set $P_\alpha$ shrinks to the set of optimal committed price distributions:

**Lemma 1.** Consider a convergent sequence of pricing distributions $F^n$ such that the corresponding profit levels $\pi^n$ converge to $\pi^*$. Then $F^n$ converges to an optimal committed price distribution.\footnote{The reason we specify the crazy type to play a strategy close to the optimal committed strategy rather than the optimal strategy itself is that the optimal strategy has a “knife-edge” feature: every sale price is charged with just enough probability for the consumer to buy at the next highest price. With such knife-edge strategies, in order for the consumer to make a plan to buy, it is not sufficient for her to be almost convinced that the monopolist is the crazy type.}

**Proof.** Note: this proof relies heavily on the proof of Proposition 7.

Define $z^n$ and $z^{n*}$ as in the proof of Proposition 7. Since $\pi^* > v$, for a sufficiently large $n$, if the consumer is facing the price distribution $F^n$, there is no PE in which she never buys, so that $z^n > 0$. As in Proposition 7, $F^n \rightarrow \Pi^*$ has a subsequence in which $z^n$ and $z^{n*}$ are constant, and the...
locations of all the atoms and their weights converge. Abusing notation somewhat, suppose $F^n$ has this property.

Consider the limit distribution $F$. By the arguments in Proposition 7 in any PE the consumer buys up to price $p_z$. Construct the price distribution $F'$ from $F$ by lumping all prices above $p_z$ up to $p_z^*$ into one atom $p'$ at the average of these prices, and moving all prices above $p_z^*$ (if any) high enough so that the consumer does not buy at these prices. Note that when facing $F$, the consumer prefers the strategy of buying up to $p_z^*$ to buying up to $p_z$. Then, when facing $F'$, in PPE the consumer buys up to $p'$, and hence the firm earns profits $\pi^*$. Hence, $F'$ must be an optimal price distribution. Therefore, in $F$ there cannot be any price atoms above $p_z^*$. Furthermore, if there was more than one atom above $p_z$, as in the proof of Proposition 7 we could create $F'$ from $F$ by replacing prices above $p_z$ by a single atom slightly above the average, earning profits strictly greater than $\pi^*$. Since this is impossible, $F$ must have a single atom above $p_z$, so that it is an optimal price distribution.

In this sense, strategies that earn profits close to $\pi^*$ are close to an optimal committed price distribution. Let $\Pi_t$ be the monopolist’s strategy in period $t$, and define $h_T(a) = (\sum_0^T \delta^t I(\Pi_t \in \mathcal{P}_a))/(1 - \delta^{T+1})$. The variable $h_T$ is a discounted count of how often the monopolist plays within $\mathcal{P}_a$.

**Proposition 6.** For any $\gamma > 0$ and $a$ satisfying $0 < \epsilon < a$, if $\delta$ is sufficiently close to 1, the monopolist earns expected discounted average profits of at least $\pi^* - a$ in any Nash equilibrium. For any $\phi > 0$, $\gamma > 0$, $a > 0$, and $w < 1$, if $\epsilon$ is sufficiently close to zero and $\delta$ is sufficiently close to 1, then in any Nash equilibrium $h_T(a) > w$ for $T$ sufficiently large. A Nash equilibrium in which the monopolist plays $\Pi_t^*$ in all periods exists.

**Proof.** Let $b_t = \Pr[\Pi_t = \Pi_t^* | \Pi_{\tau} = \Pi_{\tau}^*$ for all $\tau \leq t]$ be the consumer’s belief that if the monopolist played $\Pi_t^*$ in every period up to $t - 1$, it will do so again in period $t$. By Lemma 1 of Fudenberg and Levine (1989, page 764), $b_t \rightarrow 1$ as $t \rightarrow \infty$, and there is a minimum bound on the speed of convergence that depends only $\gamma$.

Consider first the strategy of playing $\Pi_t^*$ in every period. Note that if the consumer believes with sufficiently large probability that the monopolist will keep playing $\Pi_t^*$, her PPE is to buy at
all prices. Since by the above argument the consumer will eventually believe so, eventually she will buy at all prices. This implies that if \( \delta \) is sufficiently close to 1, the monopolist will earn an expected discounted profit of at least \( \pi^* - 2\epsilon \).

Now, it is clear that for a sufficiently small \( \epsilon > 0 \), to earn this expected discounted average profit, the monopolist must actually choose a pricing strategy in \( P_a \) at least a proportion \( w \) of the time. Either the monopolist chooses \( \Pi^*_\epsilon \) in every period, in which case the statement is trivially true, or in some period it chooses another strategy. Once it does so, the consumer knows that the monopolist is rational. In equilibrium, then, the consumer correctly anticipates the monopolist’s strategy. Hence, in at most one period can the monopolist earn expected profits greater than \( \pi^* \) (but even in that period, less than \( \bar{p} \)), and it cannot lose more than \( a \) in too many periods.

Finally, we argue that playing \( \Pi^*_\epsilon \) in every period is a Nash equilibrium for sufficiently high discount factors. It is easy to show that once the monopolist reveals itself to be a rational type, it is a continuation equilibrium for all subsequent consumers to stay away in every period. The result trivially follows.  

\[ \square \]

### B Proofs

First, we introduce some notation we will use throughout our proofs. For any market price distribution \( \Pi \), let \( p_1 \) be the lowest price, \( p_2 \) the second lowest, etc. Let \( q_l \) be the probability that \( p_l \) is charged. For notational convenience, let \( Q_l = \sum_{l'=1}^{l} q_{l'} \) and \( P_l = E[p_{l'} | l' \leq l] \).

For future reference, observe that the ex-ante expected utility when facing a market price distribution \( \Pi \) and buying at all prices less or equal to \( p_l \) is:

\[
EU(p_l; \Pi) = Q_l v - Q_l P_l - \eta(\lambda - 1)Q_l(1 - Q_l)v - \eta(\lambda - 1)Q_l(1 - Q_l)P_l - \eta(\lambda - 1)\sum_{l'=1}^{l} \sum_{l''=1}^{l'} q_{l'} q_{l''} (p_{l'} - p_{l''}).
\]

Finally, buying for all prices less or equal to \( p_l \) is a personal equilibrium if the given that the consumer expects to buy for all prices less than or equal to \( p_l \), she prefers to buy at price \( p_l \) and prefers not to buy at \( p_{l+1} \), where we set \( p_{l+1} = \infty \) if \( p_l \) is the highest price in the market price...
distribution. Hence, \( p_l \) is a personal equilibrium cutoff if and only if

\[
p_l \leq \frac{1 + \eta(1 - Q_l) + \eta \lambda Q_l}{1 + \eta \lambda} v + \frac{\eta(\lambda - 1)}{1 + \eta \lambda} Q_l P_l < p_{l+1}.
\]  \( (8) \)

### B.1 Discrete Version of Proposition 1

To establish Proposition 1, which is stated for the limit-optimal distribution, we begin by stating and proving a version of the proposition for \( \Delta > 0 \) (that is, not in the limit). To state the proposition as well as later results, we define

\[
q^*(p) = \frac{A \Delta}{(v + p)},
\]

where \( A \equiv (1 + \eta \lambda)/(\eta(\lambda - 1)) \).

We first prove the following proposition:

**Proposition 7.** For any \( \eta > 0, \lambda > 1 \), and \( \Delta \) satisfying \( 0 < \Delta < v - p \), if the firm can profitably sell to the consumer, then a profit-maximizing price distribution exists, and induces purchase with probability one. In addition, for any profit-maximizing price distribution, there exists a \( z > 0 \) such that the distribution has atoms at \( p_1, p_2, p_3, \ldots, p_z \), and \( p_{\ast} > p_z \), where \( p - 2\Delta < p_1 \leq p \), and for \( 2 \leq l \leq z \), \( p_l - p_{l-1} < 2\Delta \). For \( l < z \), the weight on atom \( p_l \) is \( q_l = \frac{A(p_{l+1} - p_l)}{(v + p_l)} \), the weight on atom \( p_z \) is \( q_z < \frac{2A\Delta}{(v + p_z)} \), and the weight on atom \( p_{\ast} \) is the complementary probability \( 1 - \sum_{l=1}^{z} q_l \).

It is useful to first outline the broad steps in our proof. There are two major steps, and several substeps.

Step I. We show that any profit-maximizing price distribution has the properties identified in the proposition. We do so by showing that for any other distribution, there is a distribution satisfying these properties that yields higher profits.

Step II. We show that among price distributions satisfying the properties of the proposition, a profit-maximizing price distribution exists.

---

\[30\] Proposition 7 is stated for any \( \Delta > 0 \). For sufficiently small \( \Delta > 0 \), we know somewhat more about the structure of the optimal price distribution. In particular, using the notation of the proposition, in that case \( p_1 = p \), \( p_{l+1} - p_l = \Delta \) for any \( l < z \), and \( q_z \leq A\Delta/(v + p_z) \).
Proof. Step I. Let $Q_z \geq 0$ the probability such that in any PE, the consumer buys the product with at least probability $z$. Furthermore, let $Q_{z^*} \geq Q_z$ be the probability with which she buys the product. Let the corresponding cutoff prices (defined as the highest atoms on the price distribution at which the consumer buys) be $p_z$ and $p_{z^*}$, respectively, and let $F$ be the optimal price distribution.

First, we show that there must be a single atom on the interval $(p_z, p_{z^*}]$ because otherwise, the monopolist could replace the stochastic prices with a single higher average price without eliminating the PPE, increasing revenues. To see this formally, suppose by contradiction that the optimal price distribution $F$ puts positive weight on more than one atom in $(p_z, p_{z^*}]$. Consider a new pricing distribution $F'$ constructed from $F$ by replacing the original prices $p_{z+1}$ through $p_{z^*}$ with the average price $p_a = \left( \sum_{l=z+1}^{z^*} p_l q_l \right) / \left( \sum_{l=z+1}^{z^*} q_l \right)$, and putting the rest of the weight on a single atom $p_a+1$ above over $p = (1 + \eta \lambda) / (1 + \eta)$. Define $Q_a$ and $P_a$ correspondingly to the notation above. Then, by construction $Q_{z^*} = Q_a$ and $Q_{z^*} P_{z^*} = Q_a P_a$. Using that for the market price distribution $F$, $p_{z^*}$ satisfies equation [8] one has

$$p_a < p_{z^*} \leq \frac{1 + \eta (1 - Q_{z^*}) + \eta \lambda Q_{z^*}}{1 + \eta \lambda} v + \frac{\eta (\lambda - 1)}{1 + \eta \lambda} Q_{z^*} P_{z^*} \leq p_{z^*+1},$$

and since $p_a < p_{z^*}$, this implies

$$p_a < \frac{1 + \eta (1 - Q_a) + \eta \lambda Q_a}{1 + \eta \lambda} v + \frac{\eta (\lambda - 1)}{1 + \eta \lambda} Q_a P_a < p_{a+1}.$$

Hence, when facing the price distribution $F'$ buying up to the price $p_a$ is a personal equilibrium. Furthermore, it is easy to show using Equation [7] that $EU(p_{z^*}; F) < EU(p_a; F')$, and by construction, $EU(p_z; F) = EU(p_z; F')$ for any $z < z^*$. Thus buying for any price less or equal to $p_a$ is the PPE strategy of the consumer when facing $F'$. Continuity of both ex-ante and ex-post utility with respect to $p_a$ implies that if the monopolist increases $p_a$ slightly the PPE still involves the consumer buying for all prices less than or equal to $p_a$. This increases profits, a contradiction.

Second, we show by contradiction that $Q_{z^*} = 1$. Suppose $Q_{z^*} < 1$. If the monopolist can profitably sell to the consumer, it must make a profit at the highest price $p_{z^*}$ at which the consumer buys in PPE. Now consider the distribution $F'$ constructed from $F$ by moving the probability weight $1 - Q_{z^*}$ from the prices above $p_{z^*}$ to $p_{z^*}$. We show that the consumer buys for all prices in the PPE for $F'$, and, hence, this change increases profits, yielding a contradiction. If $z = z^*$, it follows from
Equation 8 that buying at all prices is the unique PE with $F'$. If $z^* > z$, the above implies that $z^* = z + 1$. In addition, it follows from Equation 8 that buying at all prices is a PE after the price change. Now using Equation 7 and the fact that with price distribution $F$ the consumer prefers the PE in which she buys up to $p_{z^*}$, one has

$$EU(p_z; F) = Q_z v - Q_z P_z - \eta(\lambda - 1)Q_z(1 - Q_z)v$$

$$- \eta(\lambda - 1)Q_z(1 - Q_z)P_z - \eta(\lambda - 1)\sum_{l=1}^{z} \sum_{l' \leq 1} q_{l'} q_{l'} (p_{l'} - p_{l''})$$

$$\leq Q_z v - Q_z P_z + q_{z^*}(v - p_{z^*}) - \eta(\lambda - 1)(Q_z + q_{z^*})(1 - Q_z - q_{z^*})v$$

$$- \eta(\lambda - 1)(1 - Q_z - q_{z^*})(Q_z P_z + q_{z^*}p_{z^*})$$

$$- \eta(\lambda - 1)\left(\sum_{l' \leq 1} \sum_{l'' = 1} q_{l'} q_{l''} (p_{l'} - p_{l''}) + q_{z^*} \sum_{l = 1}^{z} q_l (p_{z^*} - p_l)\right)$$

$$= EU(p_{z^*}; F).$$

Rewriting using that

$$q_{z^*} \sum_{l=1}^{z^*} q_l (p_{z^*} - p_l) = q_{z^*}(Q_z p_{z^*} - Q_z P_z)$$

gives

$$0 \leq q_{z^*}(v - p_{z^*}) - \eta(\lambda - 1)\left( (q_{z^*}(1 - Q_z) - q_{z^*}Q_z - q_{z^*}^2)v + (1 - Q_z - q_{z^*})q_{z^*} p_{z^*} - q_{z^*} Q_z p_{z^*}\right).$$

Dividing by $q_{z^*}$, one has

$$0 \leq v - p_{z^*} - \eta(\lambda - 1)\left( ((1 - Q_z) - Q_z - q_{z^*})v + (1 - Q_z - q_{z^*})p_{z^*} - Q_z p_{z^*}\right).$$

As the right hand-side is increasing in $q_{z^*}$ and we construct $F'$ by moving the probability weight $1 - Q_z$ from the prices above $p_{z^*}$ to $p_{z^*}$, which increases $q_{z^*}$, it follows that $EU(p_z; F') \leq EU(p_{z^*}; F')$. This completes the proof that $Q_z = 1$.

Summarizing, so far we have shown that the optimal price distribution has the following structure. The monopolist charges the prices $p_1$ through $p_z$ with a total probability of $Q_z$, and the price $p_{z^*}$ with probability $1 - Q_z$, where either $z^* = z$ or $z^* = z + 1$. In addition, if $z^* = z$, there is exactly one PE, and if $z^* = z + 1$, there are exactly two PE: one in which the consumer buys up to price $p_z$, and one in which she buys at all prices. Finally, in the PPE the consumer buys at all
prices. Our next goal is to show that in the optimal price distribution, we have $0 < Q_z < 1$, so that $z^* = z + 1$ and $z > 0$. We establish this by showing that the monopolist can earn greater revenue with $z^* = z + 1$ and $z > 0$ than with either $z = 0$ or $z^* = z$.

First, consider $z = 0$. In that case, the monopolist charges a single deterministic price, and we have already shown in the text that the optimal deterministic price is $v$.

Now consider the case $z > 0$. Note that if $z^* = z + 1$, then for the consumer to be willing to buy at all prices, it must both be a PE to buy up to price $p_z$, and this strategy must be preferred to the PE of buying only up to price $p_z$. By Equations 8 and 9, the highest $p_z$ at which this holds is

$$p_z^* = \min \left\{ \frac{v + \eta(\lambda-1)Q_zP_z}{1 - \frac{\eta(\lambda-1)}{1+\eta\lambda}q_z}, v + \frac{2\eta(\lambda-1)Q_zP_z}{1 + \eta(\lambda-1)Q_z} \right\}$$

(10)

Notice that that holding $Q_z$ fixed (which also fixes $q_z^* = 1 - Q_z$), $p_z^*$ is increasing in $Q_zP_z$. Hence, whether or not $z^* = z$ or $z^* = z + 1$, in order to maximize profits the monopolist must maximize $Q_zP_z$ subject to the constraint that the consumer buys with probability $Q_z$ in any PE.

We next consider the implications of this maximization problem.

Notice that for any price $p_l < p_z$ on the support of the distribution, we show by contradiction that it is optimal to charge $p_l$ with the lowest possible probability such that the consumer is just willing to buy at the next price if she had been expecting to buy at prices up to $p_l$. Suppose this is not the case, and consider shifting a little bit of weight from $p_l$ to $p_{l+1}$. For a sufficiently small shifted weight, Equation 8 implies that it will still be the case that in any PE the consumer buys at all prices up to $p_z$.

We now solve for the weight the monopolist must put on each price for the the above property to hold for all $l < z$. That the consumer is just willing to buy at price $p_l$ if she had been expecting to buy at prices up to $p_{l-1}$ is equivalent to

$$v - p_l + \eta(1 - Q_{l-1})v - \eta\lambda(1 - Q_{l-1})p_l - \eta\lambda Q_{l-1}(p_l - P_{l-1}) \geq -\eta\lambda Q_{l-1}v + \eta Q_{l-1}P_{l-1},$$

or

$$(1 + \eta + \eta(\lambda - 1)Q_{l-1})v - (1 + \eta\lambda)p_l + \eta(\lambda - 1)Q_{l-1}P_{l-1} = 0.$$
The corresponding equation for the consumer to just be willing to buy at price \( p_{l+1} \) is

\[
(1 + \eta + \eta(\lambda - 1)Q_l)v - (1 + \eta\lambda)p_{l+1} + \eta(\lambda - 1)Q_lP_l = 0.
\]

Subtracting the latter equation from the former one and rearranging yields

\[
q_l = \frac{(1 + \eta\lambda)(p_{l+1} - p_l)}{\eta(\lambda - 1)(v + p_l)} = \frac{A(p_{l+1} - p_l)}{v + p_l}.
\]

This completes the claim in the proposition regarding the weights \( q_l \) for \( l < z \).

Next, we establish that \( \text{Pr}_F(p_z) < 2A\Delta/(v + p_z) \). Suppose by contradiction that \( \text{Pr}_F(p_z) \geq 2A\Delta/(v + p_z) \). Then, if the monopolist set \( p_{z^*} = p_z + 2\Delta \), it would be a unique PE for the consumer to buy at all prices. Hence, the optimal price distribution must have \( p_{z^*} > p_z + 2\Delta \). Hence, the monopolist could construct a new distribution \( F' \) from \( F \) in the following way. Let \( z' = z + 1 \), \( z'^{*} = z^{*} + 1 \), with the distribution \( F' \) created from \( F \) by shifting up the weight \( \text{Pr}_F(p_z) - A\Delta/(v + p_z) \) from \( p_z \) to \( p_{z+1} = p_z + \Delta \). Then, by the above calculation, with \( F' \) the consumer buys up to \( p_{z+1} \) in any PE. Since \( Q'_zp'_z > Q_zp_z \), this contradicts that \( Q_zp_z \) maximizes profits subject to the constraint that the consumer buys with probability \( Q_z \) in any PE.

Now we show that up to \( p_z \) the atoms of the optimal price distribution are spaced at intervals of less than \( 2\Delta \). Suppose by contradiction that this is not the case for the optimal price distribution \( F \), so that for some \( l \leq z - 1 \), \( p_{l+1} - p_l \geq 2\Delta \). We construct the distribution \( F' \) from \( F \) in the following way. We let \( z' = z + 1 \) and \( z'^{*} = z^{*} + 1 \), we put an extra atom at \( p_l + \Delta \), and let \( q'_l = A\Delta/(v + p_l) \) and \( q'_{l+1} = q_l - A\Delta/(v + p_l) \), with the weights and positions of the other atoms remaining the same. Since \( q'_{l+1} = A(p'_{l+2} - p'_{l+1})/(v + p_l) > A(p'_{l+2} - p'_{l+1})/(v + p'_{l+1}) \), this maintains the property that in any PE the consumer buys at all prices up to \( p_{z}(= p'_{z+1}) \). And since \( Q'_zp'_z > Q_zp_z \), this contradicts that \( Q_zp_z \) maximizes profits subject to the constraint that the consumer buys with probability \( Q_z \) in any PE.

Next, we show that \( p - 2\Delta < p_1 \leq p \). Clearly, if \( p_1 > p \), there is a PE in which the consumer does not buy. We are left to show that \( p_1 > p - 2\Delta \). Suppose otherwise. Then, since \( p_2 - p_1 < 2\Delta \), we must have \( p_2 < p \). Now we construct the price distribution \( F' \) from \( F \) by moving the atom at \( p_1 \) to \( p_2 \). This ensures that the consumer buys for all prices up to \( p_z \) in any PE, and has \( Q'_zp'_z > Q_zp_z \), a contradiction.
We now establish that if $\Delta < v - p$, the firm charges at least two prices with positive probability, so that $z > 0$. Recall that the optimal deterministic price is $v$. To prove that the firm charges at least two prices with positive probability, we construct a hybrid distribution with which the monopolist earns expected revenue greater than $v$. Consider the distribution that puts weight $\epsilon > 0$ on $p$ and weight $1 - \epsilon$ on $p_z*$ as defined in Equation 10. Note that for a sufficiently small $\epsilon$, the minimum in Equation 10 is determined by the second argument in the minimum function. Hence, with this pricing distribution the firm’s revenue is:

$$\int p h(p) dp = \left[ \frac{2 + \eta + \eta \lambda}{\eta (\lambda - 1)} \left( \exp \left( \frac{\eta (\lambda - 1)}{1 + \eta \lambda} \right) - 1 \right) - 1 \right] \cdot v < v. \quad (12)$$

First, we calculate $p_{max}$, which solves

$$\int_{p}^{p_{max}} \frac{1}{v + p} dp = 1.$$
This gives
\[ \ln \left( \frac{v + p_{\text{max}}}{v + p} \right) = \frac{\eta(\lambda - 1)}{1 + \eta \lambda}, \]
or
\[ \frac{p_{\text{max}}}{v} = \exp \left( \frac{\eta(\lambda - 1)}{1 + \eta \lambda} \right) \left( 1 + \frac{1 + \eta}{1 + \eta \lambda} \right) - 1. \]

Now the expected revenue with price distribution \( h(\cdot) \) is
\[ \frac{1 + \eta \lambda}{\eta(\lambda - 1)} \int_{p}^{p_{\text{max}}} \frac{p}{v + p} \, dp = \frac{1 + \eta \lambda}{\eta(\lambda - 1)} \int_{p}^{p_{\text{max}}} \left( 1 - \frac{v}{v + p} \right) \, dp = \frac{1 + \eta \lambda}{\eta(\lambda - 1)} (p_{\text{max}} - p) - v. \]

Plugging in for \( p_{\text{max}} \) and rearranging gives the expression in Equation 12.

**Step II.** Suppose by contradiction that a profit-maximizing pricing distribution does not exist. Then, since the firm’s profits are bounded, there must be a sequence of price distributions \( F^n \) such that the corresponding profits converge to the supremum profit level \( \pi^* \). By the logic of Step I, for any pricing distribution there is a corresponding pricing distribution with at least as high profits that satisfies the properties of the proposition, and for which the highest price is given by Equation 10. Hence, we can choose \( F^n \) so that it satisfies these properties.

Define by \( z^n \) and \( z^{n*} \) for each \( F^n \) as above. Since pricing atoms must be at least \( \Delta \) apart, and the consumer does not buy for any price about \( \bar{p} \), \( z^n \) and \( z^{n*} \) both come from a finite set. Therefore, \( F^n \) must have a subsequence for which \( z^n \) and \( z^{n*} \) is constant. With slight abuse of notation, we assume that \( F^n \) already has this property. Then, by the diagonal method, it is easy to show that \( F^n \) has a subsequence in which the locations of all atoms and all their weights converge. With another slight abuse of notation, we assume that \( F^n \) already has this property.

Now consider the limiting distribution of the sequence \( F^n \), \( F \). By construction, in any PE the consumer buys for any price up to \( p_x \). In addition, by Equation 10 which is continuous in \( p_t \) and \( q_t \), in PPE the consumer is willing to buy also at \( p_x \). Hence, when facing \( F \), the PPE is for the consumer to buy at all prices, so that the firm achieves profit level \( \pi^* \)—a contradiction.

**B.2 Proof of Propositions 1, 2, and 3**

**Proof of Proposition 1**
Consider a sequence $\Delta_n \to 0$ such that a sequence of corresponding optimal pricing distributions $F^n$ converge in distribution. Define $z^n, p^n_l, q^n_l,$ and $Q^n_{z^n}$ analogously to Proposition 7. Assume first that $Q^n_{z^n}$ converges to some $s$; we will establish this below.

Trivially, as $\Delta$ decreases the optimal profits must weakly increase since the firm could always choose the same distribution as it did for a higher value of $\Delta$. Also, the profits the monopolist can earn are bounded, so that there is a limiting profit strictly greater than $v$. By the proof of Proposition 7, if we had $s = 0$, then the limiting profit would be $v$, and if we had $s = 1$, the limiting profits would be less than $v$. Hence, we can conclude that $0 < s < 1$.

As in Proposition 7, consider the distribution on $[p, p_{\text{max}}]$ with density

$$h(p) = \frac{1 + \eta \lambda}{\eta(\lambda - 1)(v + p)} = \frac{A}{v + p}.$$ 

Let the corresponding cumulative distribution function be $H$, and define $p_{\text{max}}(s)$ so that $H(p_{\text{max}}(s)) = s$. We now establish that for $x \leq p_{\text{max}}(s)$, $F^n(x) \to H(x)$ as $n \to \infty$; that is, in that part of the real line $F^n$ converges in distribution $H$.

Since $p - 2\Delta_n < p^n_l \leq p$, we have $p^n_l \to p$. We prove that $p^n_{z^n} \to p_{\text{max}}(s)$. We have

$$Q^n_{z^n} = \sum_{l=1}^{z^n} q^n_l = q^n_{z^n} + A \sum_{l=1}^{z^n-1} \frac{p^n_{l+1} - p^n_l}{v + p^n_l} = q^n_{z^n} + A \sum_{l=1}^{z^n-1} \left[ \int_{p^n_l}^{p^n_{l+1}} \frac{1}{v + p} dp + \int_{p^n_l}^{p^n_{l+1}} \left( \frac{1}{v + p^n_l} - \frac{1}{v + p} \right) dp \right]$$

(13)

We work on the sum of the underbraced term:

$$\sum_{l=1}^{z^n-1} \int_{p^n_l}^{p^n_{l+1}} \left( \frac{1}{v + p^n_l} - \frac{1}{v + p} \right) dp = \sum_{l=1}^{z^n-1} \int_{p^n_l}^{p^n_{l+1}} \frac{p - p^n_l}{(v + p^n_l)(v + p)} dp.$$ 

Notice that this is positive and (since $p^n_{l+1} - p^n_l < 2\Delta_n$) it is less than

$$\sum_{l=1}^{z^n-1} \int_{p^n_l}^{p^n_{l+1}} \frac{2\Delta_n}{(v + p^n_l)(v + p)} dp < \sum_{l=1}^{z^n-1} 2\frac{(p^n_{l+1} - p^n_l)\Delta_n}{v^2} = \frac{2(v_{z^n}^n - p^n_l)\Delta_n}{v^2},$$

which approaches zero as $n \to \infty$. Taking the limit of Equation 13, plugging in that the sum of the underbraced terms approaches zero, and using that $q^n_{z^n} \to 0$ as $n \to \infty$, we get

$$s = \lim_{n \to \infty} A \int_{p^n_l}^{p^n_{z^n}} \frac{1}{v + p} dp = \lim_{n \to \infty} A \int_{p}^{p^n_{z^n}} \frac{1}{v + p} dp.$$ 

48
This implies that \( p_{zn}^n \to p_{\text{max}}(s) \) as \( n \to \infty \).

Next, we show that for a sufficiently large \( n \), we have \( p_{zn+1}^n > p_{\text{max}}(s) \). We know that \( p_{zn}^n \) satisfies the condition that if the consumer expected to buy up to price \( p_{zn-1}^n \), she would just be indifferent to buying at \( p_{zn}^n \). This is equivalent to
\[
p_{zn}^n = \frac{(1 + \eta + \eta(\lambda - 1)Q_{zn-1}^n)v + \eta(\lambda - 1)Q_{zn-1}^n P_{zn-1}^n}{1 + \eta \lambda} \leq \frac{(1 + \eta + \eta(\lambda - 1)Q_{zn}^n)v + \eta(\lambda - 1)Q_{zn}^n P_{zn}^n}{1 + \eta \lambda}
\]

Given that \( p_{zn}^n \to p_{\text{max}}(s) \) and \( Q_{zn}^n \to s < 1 \), this and Equation 10 imply that for a sufficiently large \( n \), we have \( p_{zn+1}^n > p_{\text{max}}(s) \).

Clearly, for any \( x \leq p \), \( H(x) = \lim_{n \to \infty} F^n(x) = 0 \). Now take any \( x \) satisfying \( p < x < p_{\text{max}}(s) \). So long as \( p_{zn}^n > x \), which holds for \( n \) sufficiently large, we have
\[
F^n(x) = \sum_{l \cdot p_l^n \leq x} \frac{q_l^n}{v + p_l^n} = A \sum_{l \cdot p_l^n \leq x} \left[ \int_{p_l^n}^{p_{l+1}^n} \frac{1}{v + p} dp + \int_{p_l^n}^{p_{l+1}^n} \left( \frac{1}{v + p} - \frac{1}{v + p_l^n} \right) dp \right].
\]

(14)

By the same argument as above, the sum of the underbraced term approaches zero as \( n \to \infty \), and we must have \( \max_{l} \{ p_l^n | p_l^n \leq x \} \to x \) as \( n \to \infty \). Hence, taking the limit of Equation 14, we have
\[
\lim_{n \to \infty} F^n(x) = A \int_{p}^{x} \frac{1}{v + p} dp = H(x).
\]

Finally, since for \( n \) sufficiently large \( p_{zn+1}^n > p_{\text{max}}(s) \), \( \lim_{n \to \infty} \Pr_{F^n}(p_{\text{max}}(s)) = 0 \). This completes the proof that for \( x \leq p_{\text{max}}(s) \), \( F^n(x) \to H(x) \) as \( n \to \infty \).

Next, notice that in order for \( F^n \) to converge in distribution, the sequence \( p_{zn+1}^n \) must converge. Let the limit be \( p \). Applying Equation 10, \( p > p_{\text{max}}(s) \). We have shown that the limiting distribution has the properties in the proposition.

To conclude the proof, it remains to show that \( Q_{zn}^n \) converges. Suppose by contradiction that it does not. Then, the sequence \( F^n \) must have two subsequences \( F^{n_1} \) and \( F^{n_2} \) such that \( Q_{zn_1}^{n_1} \) and \( Q_{zn_2}^{n_2} \) both converge, but to different limits \( s_1 \) and \( s_2 \), respectively. Then, the above arguments imply that \( F^{n_1} \) and \( F^{n_2} \) converge in distribution to different distributions: the limit of \( F^{n_1} \) is distributed continuously on \([p, p_{\text{max}}(s_1)]\) and has an isolated atom, while the limit of \( F^{n_2} \) is distributed continuously on \([p, p_{\text{max}}(s_2)]\) and has an isolated atom. But this means that the sequence \( F^n \) does not converge in distribution, a contradiction.

\( \square \)
Proof of Proposition 2. From the proof of Proposition 7, for $\Delta < v - p$ the consumer buys the product with probability one at an expected price strictly greater than $v$. Hence, her consumption utility is negative. Furthermore, in any PE expected gain-loss utility is non-positive. If she follows through a plan of never buying, both her consumption utility and her gain-loss utility are zero. □

Proof of Proposition 3. Note: this proof relies heavily on the proof of Proposition 1.

We first state a version of our result for $\Delta > 0$. As in the proposition, let $p' = (1 + \eta + \eta(\lambda - 1)\phi/p)v/(1 + \eta \lambda)$.

Lemma 2. Fix any $\phi > 0, \eta > 0, \lambda > 1$. Then, for a sufficiently small $\Delta > 0$, if the firm can profitably sell to the consumer, then a profit-maximizing price distribution exists, and induces purchase with probability one. Furthermore, there is a unique cutoff $\phi^c(\Delta) < p$ such that for $\phi > \phi^c(\Delta)$, the unique price distribution puts probability one on $v - \phi$; and for $\phi < \phi^c(\Delta)$, there exists a $z > 0$ such that the distribution has atoms at $p_0, p_1, p_2, p_3, \ldots, p_z$, and $p_z^* > p_z$, where $p_0 = 0$ and $q_0 = \phi/p$, $p' - 2\Delta < p_1 \leq p'$, and for $2 \leq l \leq z, p_l - p_{l-1} < 2\Delta$. For $l < z$, the weight on atom $p_l$ is $q_l = A(p_{l+1} - p_l)/(v + p_l)$, the weight on atom $p_z$ is $q_z < 2A\Delta/(v + p_z)$, and the weight on atom $p_z^*$ is the complementary probability $1 - \sum_{l=0}^{z} q_l$.

Proof. By essentially the same proof as in Proposition 7, the profit-maximizing price distribution induces purchase with probability one, and the highest revenue the monopolist can earn with a deterministic price distribution is $v - \phi$. Next, we establish some properties that an optimal price distribution must have for it to earn revenue strictly greater than $v - \phi$. Then, we complete the proof by showing that there is a cutoff such that the latter type of distribution earns revenues less than $v - \phi$ for $\phi$ above the cutoff, while it earns revenues greater than $v - \phi$ for $\phi$ below the cutoff.

Since the price distribution features an average price greater than $v - \phi$, the consumer’s expected consumption utility in a PE in which she buys at all prices is negative. Since her gain-loss utility is less than or equal to zero, this means that she would prefer to make and follow through a plan
not to buy. Hence, in order for her to buy at all prices, a strategy of never buying must not be credible. Let \( q_0 = \Pr(p \leq p) \)—that is, \( q_0 \) is the probability that the monopolist’s price is less than \( p \). Similarly, let \( p_0 = E[p|p \leq p] \). If the consumer had expected not to buy, if she checks the price she will buy if \( p \leq p \). Hence, she will check the price if

\[
q_0(1 + \eta)v - q_0(1 + \eta\lambda)p_0 - (1 + \eta\lambda)\phi \geq 0,
\]
or

\[
q_0 \geq \frac{\phi}{p - p_0}.
\]  \( (15) \)

Now, slightly modifying the proof of Proposition 7 we consider two cases.

Case I. The consumer has a PE in which she buys only up to price \( p \). In this case, by the same logic as in the proof of Proposition 7 there is a single price above \( p \) that the monopolist charges with positive probability. Clearly, for a sufficiently small \( \Delta > 0 \) this price atom is above \( p + 2\Delta \).

Case II. The consumer has no PE in which she buys up to price \( p \). Then, by the same steps as in the proof of Proposition 7, the (truncated) price distribution above \( p \) has the same qualitative properties as with no price-discovery costs: there exists a \( z > 0 \) such that (i) the distribution has atoms at \( p_1, p_2, p_3, \ldots, p_z \), and \( p_{z*} > p_z \); (ii) for \( 2 \leq l \leq z \), \( p_l - p_{l-1} < 2\Delta \); (iii) for \( l < z \), the weight on atom \( p_l \) is \( q_l = A(p_{l+1} - p_l)/(v + p_l) \), and the weight on atom \( p_z \) is \( q_z < 2A\Delta/(v + p_z) \); and (iv) the weight on atom \( p_{z*} \) is the complementary probability \( 1 - \sum_{l=0}^{z} q_l \).

Furthermore, again by the same logic as in the proof of Proposition 7, \( p_1 \) is within \( 2\Delta \) of the highest price at which the consumer would buy if she had expected to buy up to \( p \). This means that \( p_1 > [(1 + \eta + \eta(\lambda - 1)v + \eta(\lambda - 1)q_0p_0]/(1 + \eta\lambda) - 2\Delta \). Therefore, by Equation 15 for a sufficiently small \( \Delta > 0 \) we have \( p_1 > p + 2\Delta \).

Next, we show that \( p_0 = 0 \) and \( q_0 = \phi/p \). By Equation 15 we already know that \( q_0 \geq \phi/p \). Suppose by contradiction that \( p_0 > 0 \) or \( q_0 > \phi/p \). Then, Equation 15 implies that we cannot have \( p_0 > 0 \) and \( q_0 = \phi/p \), so that we must have \( q_0 > \phi/p \). Hence, we can construct a new price distribution that puts weight \( \phi/p \) on the price of zero, and weight \( q_0 - \phi/p \) on a price of \( p + \epsilon \). For a sufficiently small \( \epsilon > 0 \), it is not credible for the consumer to buy only at price zero, so that she buys at both prices. In addition, for a sufficiently small \( \Delta > 0 \) and \( \epsilon < \Delta \), the new price distribution
does not violate the constraint that price atoms must be at least $\Delta$ apart. Furthermore, with this alternative price distribution, the expected price in this range is $(q_0 - \phi/p)(p + \epsilon) > q_0p - \phi$. By Equation 15 with the original distribution the expected price is $q_0p_0 \leq q_0p - \phi$, so that the change increases the expected price in this range. Finally, notice that with the increase in the expected price in this range, the consumer still buys at all higher prices, so that the monopolist earns greater profits overall, a contradiction.

Now suppose that we are in Case I above. Let the single price atom be $p_{reg}$. The consumer’s expected ex-ante utility if she buys only at the zero price is

$$q_0v - q_0(1 - q_0)\eta(\lambda - 1)v - \phi,$$

whereas if she buys at both prices it is

$$v - (1 - q_0)p_{reg} - q_0(1 - q_0)\eta(\lambda - 1)p_{reg} - \phi.$$

Hence, in order for her to prefer to buy at both prices it must be the case that $p_{reg} \leq v$. This means that the monopolist’s expected revenue is $(1 - \phi/up)v < v - \phi$. Therefore, with such a pricing strategy the monopolist cannot earn revenue greater than $v - \phi$. This establishes that the monopolist either chooses a price distribution from Case II above, or chooses the deterministic price $v - \phi$.

To complete the proof, we establish the existence of the cutoff $\phi^c(\Delta)$ specified in the proposition. Notice that by essentially the same proof as in Proposition 7 if $\phi$ is sufficiently small, the stochastic price distribution earns revenue greater than $v$, so that is what the monopolist will choose. Conversely, it is easy to show that for $\phi < p$ sufficiently close to $p$, the stochastic price distribution from Case II earns revenues less than $v - \phi$. To see this, notice that the consumer will never buy at a price greater than $\overline{p}$, so that the monopolist’s revenue is at most

$$\left(1 - \frac{\phi}{\overline{p}}\right)\overline{p} = v + (\overline{p} - v) - \phi - (\overline{p} - p)\frac{\phi}{\overline{p}},$$

which, since $\overline{p} < v$, is strictly less than $v - \phi$ for $\phi$ sufficiently close to $\overline{p}$.

Now, to complete the proof it is sufficient to establish that if the monopolist prefers the stochastic price distribution for some $\phi$, it strictly prefers a stochastic price distribution for $\phi' < \phi$. Take
the optimal price distribution for $\phi$, and construct a new price distribution that puts $\phi'/p$ on the price of zero and $(\phi - \phi')/u p$ on the price of $p + \epsilon$. This increases the monopolist’s revenues by more than $\phi - \phi'$, and for a sufficiently small $\epsilon > 0$, maintains the properties that atoms are at least $\Delta$ apart and that the consumer buys at higher prices. Hence, for $\phi'$ the monopolist must prefer this new distribution to the deterministic one, completing the proof.

To prove Proposition 3 from Proposition 2, we can follow the same steps as in the proof of Proposition 1.

Proof of Proposition 4. To be typed in.

Proof of Proposition 5. Note that in equilibrium firms make zero expected profits.

We first prove that there is no equilibrium in which a firm chooses a stochastic price distribution. Suppose by contradiction that firm 1 chooses a stochastic price distribution, in which it charges the prices $p_1$ through $p_I$ ordered from lowest to highest with probabilities $q_1$ through $q_I$, respectively. It is clearly optimal to associate lower costs with lower prices; hence, let the associated average costs be $c_1$ through $c_I$, respectively, which are also increasing. We suppose that a positive measure of consumers buy the product at all prices; a slight modification of the proof below covers the other case. Let $p^* = (c_L + c_H)/2$ and $W = \eta(\lambda - 1)\left(\sum_{i,i'} q_i q_{i'} |p_i - p_{i'}|\right)/2$. The variable $W$ is the expected gain-loss disutility from price variation for a consumer who buys at all prices at firm 1.

We consider the response by firm 2 of setting the deterministic price $p^*$. We show that a positive measure of consumers with $v > p^*$ strictly prefer to go to firm 2 over firm 1. This yields a contradiction because for a sufficiently small $\epsilon > 0$, if firm 1 charges $p^* + \epsilon$ with probability 1, it still attracts a positive measure of consumers and earns positive profits.

Suppose by contradiction that the measure of consumers who strictly prefer firm 2 is zero. For those who would buy the product at all prices at firm 1 to prefer firm 1, we must have $E[p] + W \leq p^*$, so that $E[p] < p^*$ and $W \leq p^* - E[p]$.

We now show by contradiction that if $d$ is sufficiently large, firm 2 cannot sell profitably to any consumer with $v \leq p^*$ by choosing a price distribution such that $E[p] < p^*$. Clearly, it is sufficient to establish this for $v = p^*$. We consider two cases. First, suppose that the expected utility of
consumer \( p^* \) expected utility is non-negative. Suppose that this consumer buys the product with probability \( q \) at an average price conditional on buying of \( p_{\text{ave}} \). Then

\[
q(p^* - p_{\text{ave}}) - \eta(\lambda - 1)q(1 - q)(p^* + p_{\text{ave}}) \geq 0,
\]

which yields

\[
q > 1 - \frac{p^* - p_{\text{ave}}}{\eta(\lambda - 1)(p^* + p_{\text{ave}})}.
\]

In order for the firm to sell profitably when the average price is \( p_{\text{ave}} \), the average cost at which it sells must be less than \( p_{\text{ave}} \). Supposing that the firm sells for cost levels on the interval \([c_L, c'_L]\), we must therefore have \((c_L + c'_L)/2 < p_{\text{ave}} \). Noting that \( c'_L = c_L + q/d \) and rearranging gives

\[
q < 1 - 2d(p^* - p_{\text{ave}}).
\] (16)

For \( d \) sufficiently large, \( q \) cannot simultaneously satisfy the above two inequalities for any \( p_{\text{ave}} < p^* \).

Second, suppose that the expected utility of consumer \( p^* \) is negative. Then, not buying cannot be a credible plan. Suppose again that consumer \( p^* \) buys with probability \( q \) at an average price conditional on buying of \( p_{\text{ave}} \). Using that \( W \leq p^* - E[p] \), we must have \( W \leq p^* - p_{\text{ave}} \). In addition, for firm 1 to sell profitably, Inequality [16] must be satisfied. We show that for \( d \) sufficiently small, there is no profitable price distribution satisfying these properties.

Note that for firm 1 to make positive profits, \( p_{\text{ave}} > c_L \) must hold. Since \( p^* - c_L = 1/(2d) \), this gives \( W \leq 1/(2d) \). Given that not buying is not a credible strategy, the same steps as in the proof of Proposition [7] show that for a sufficiently large \( d \), the optimal price distribution that induces purchase with probability \( q \) and satisfies \( W \leq 1/(2d) \) has the following properties: (i) it puts weight \( s > 0 \) on prices \( p_1, \ldots, p_z \) with the same properties as in Proposition [7] applied to \( v = p^* \); (ii) it puts weight \( q - s \) on a single atom \( p_{z^*} \); and (iii) it puts the rest of the weight on higher prices at which the consumer does not buy. Let \( p^{*}_Z = (1 + \eta)p^*/(1 + \eta\lambda) \). Notice that for \( W \) sufficiently small, it must be that \( p_1 \) through \( p_z \) are all less than \( p' = (p^* + p^{*}_Z)/2 \). Suppose \( d \) is sufficiently large to have \( c_L > p' \). Then, for firm 1 to make positive profits, we must have \( p_{z^*} > c_L > p' \). Now notice that

\[
\eta(\lambda - 1)s(1 - s)(p_{z^*} - p') < \eta(\lambda - 1)\left( \sum_{i,i'} q_i q_{i'} |p_i - p_{i'}| \right)/2 = W \leq p^* - E[p] \leq p^* - p_{\text{ave}}.
\] (17)
Now the consumer’s expected utility when she buys only at the prices $p_1$ through $p_z$ is at least $s(p^* - p') - \eta(\lambda - 1)s(1 - s)(p^* + p')$. In order for her to buy also at $p_{z*}$, her expected utility from doing so must be at least as great as her expected utility from buying only at prices $p_1$ through $p_z$. This implies

$$q(p^* - p_{ave}) - \eta(\lambda - 1)q(1 - q)(p^* + p_{ave}) > -\eta(\lambda - 1)s(1 - s)(p^* + p') > -(p^* - p_{ave})\frac{p^* + p'}{p_{z*} - p'},$$

where the last inequality follows from Inequality 17. Now notice that for any $q < 1$, if $p^* - p_{ave}$ is sufficiently small, then the above inequality implies that $q > q$. Hence, if $d$ is sufficiently large, $q > q$. Suppose, then, that $d$ is sufficiently large for us to have $q > 1/2$. Rearranging the above inequality gives

$$q > 1 - \frac{p^* - p_{ave}}{\eta(\lambda - 1)(p^* + p_{ave})} \left(1 + \frac{p^* + p'}{q(p_{z*} - p')}\right) > 1 - \frac{p^* - p_{ave}}{\eta(\lambda - 1)(p^* + p_{ave})} \left(1 + \frac{2(p^* + p')}{p_{ave} - p'}\right).$$

Now, for a sufficiently large $d$, $q$ cannot satisfy both this inequality and Inequality 16 for any $p_{ave} \geq c_L$, a contradiction. This completes the proof that firm 1 cannot profitably sell to any consumer with $v \leq p^*$.

The above implies that there is a positive measure of consumers with $v > p^*$ on whom firm 1 would not make losses if they went to firm 1. Note that since $E[p] < p^*$, such a consumer must be buying with probability less than 1 for firm 1 not to make losses. Take such a consumer, and suppose that she buys the product with probability $q$ at an average price conditional on buying of $p_{ave}$. We next show that for a sufficiently large $d$, any such consumer strictly prefers firm 2’s offer of a deterministic $p^*$. We do this by showing that if a consumer with $v > p^*$ prefers firm 1’s price distribution, then firm 1 makes losses on the consumer. That the consumer prefers firm 1’s distribution implies

$$q(v - p_{ave}) - \eta(\lambda - 1)q(1 - q)(v + p_{ave}) \geq v - p^*,$$

which in turn leads to

$$q \geq 1 - \frac{(v - p_{ave}) - (v - p^*)/q}{\eta(\lambda - 1)(v + p_{ave})} > 1 - \frac{p^* - p_{ave}}{\eta(\lambda - 1)(p^* + p_{ave})}.$$  

This implies that for a sufficiently large $d$, $q$ violates Inequality 16 with the inequality going strictly the other way, for any $p_{ave}$ satisfying $c_L \leq p_{ave} < p^*$, so that firm 1 makes expected losses on this
consumer. for any $p_{ave} \geq c_L$. This completes the proof that there is no equilibrium in which one firm chooses a stochastic price distribution.

Clearly, if both firms choose deterministic price distributions, their price must equal $p^*$. To complete the proof, we show that for a sufficiently large $d$, this is indeed an equilibrium. Suppose firm 2 sets the deterministic price $p^*$. If firm 1 does the same, it gets zero expected profits. To show that firm 1 has no profitable deviation, we show that it cannot attract any consumer from firm 2 and make strictly positive profits on the consumer. Consider first consumers with $v \leq p^*$. These consumers get an expected utility of zero from going to firm 2. We have shown above that firm 1 cannot make positive profits on a consumer with $v < p^*$ such that the consumer’s expected utility is non-negative. Hence, firm 1 cannot profitably attract these consumers.

Now consider consumers with $v > p^*$. Clearly, firm 1 cannot profitably attract these consumers in a way that leads them to buy with probability one. And we have shown above that firm 1 cannot profitably attract these consumers and have them buy with probability less than one. This completes the proof. $\square$